

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 22, 2006, 01:28:24 ; Search time 189 Seconds  
(without alignments)  
1011.269 Million cell updates/sec

Title: US-10-717-619-2  
Perfect score: 2316  
Sequence: 1 MTSQGNKRTTKKEFGDLRFQ.....QPLPLARLLLTQYSSQALHE 435

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_21.\*  
1: Geneseqp1980s.\*  
2: Geneseqp1990s.\*  
3: Geneseqp2000s.\*  
4: Geneseqp2001s.\*  
5: Geneseqp2002s.\*  
6: Geneseqp2003as.\*  
7: Geneseqp2003bs.\*  
8: Geneseqp2004s.\*  
9: Geneseqp2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2316	100.0	435	4 AAB81070	Aab81070 Murine ma
2	1316.5	56.8	428	5 AAE22609	Aae22609 Human MIS
3	1316.5	56.8	428	7 ADB65190	Adb65190 Human pro
4	1316.5	56.8	443	5 AAE22608	Aae22608 Human MIS
5	1132	48.9	376	4 AAB81071	Aab81071 Human mas
6	1109.5	47.9	353	5 AAE22610	Aae22610 Human MIS
7	398	17.2	533	7 ADP65191	Adp65191 Human lym
8	398	17.2	533	8 ADQ19792	Adq19792 Human sof
9	398	17.2	533	8 ADP23847	Adp23847 PRO poly
10	398	17.2	533	8 ADP23845	Adp23845 PRO poly
11	398	17.2	533	8 ADT07538	Adt07538 Human col
12	337	14.6	224	8 ADP45459	Adp45459 Human col
13	337	14.6	225	8 ADP45461	Adp45461 Human col
14	337	14.6	230	8 ADP45460	Adp45460 Human col
15	326	14.1	456	4 AAB93406	Aab93406 Human pro
16	319	13.8	449	2 AAW59866	Aaw59866 Amino aci
17	319	13.8	456	2 AAW59865	Aaw59865 Amino aci
18	319	13.8	456	9 ADZ26491	Adz26491 Human BLI
19	287.5	12.4	457	2 AAW59867	Aaw59867 Amino aci
20	260	11.2	95	2 AAW70587	Aaw70587 Human SLP
21	213	9.2	84	3 AAG03152	Aag03152 Human sec
22	212	9.2	46	4 AAM18973	Aam18973 Peptide #
23	212	9.2	46	4 ABB38147	Abb38147 Peptide #
24	212	9.2	46	4 AAM31575	Aam31575 Peptide #

25	212	9.2	46	4 ABB23354	Abb23354 Protein #
26	212	9.2	46	4 AAM71296	Aam71296 Human bon
27	212	9.2	46	4 AAM58783	Aam58783 Human bra
28	212	9.2	46	4 ABG53007	Abg53007 Human liv
29	212	9.2	46	5 ABG41096	Abg41096 Human pep
30	185.5	8.0	503	4 ABG22854	Abg22854 Novel hum
31	185.5	8.0	503	7 ADC32915	Adc32915 Human nov
32	165.5	7.1	474	2 AAM15253	Aam15253 Human bra
33	165.5	7.1	594	2 AAM15256	Aam15256 Human bra
34	165.5	7.1	594	9 ADX05533	Adx05533 Cyclin-de
35	165.5	7.1	594	9 ADZ28263	Adz28263 Human SHC
36	161.5	7.0	344	7 ADE14373	Ade14373 Human int
37	150.5	6.5	743	4 AAM79738	Aam79738 Human pro
38	146	6.3	728	7 ADD46841	Add46841 Human pro
39	146	6.3	728	7 ADE56294	Ade56294 Human pro
40	146	6.3	728	8 ADJ66562	Adj66562 P13 kinas
41	144.5	6.2	474	2 AAM39087	Aam39087 Human SHC
42	144.5	6.2	474	2 AAM39089	Aam39089 Human SHC
43	143.5	6.2	722	7 ADD46839	Add46839 Rat Prote
44	143.5	6.2	722	7 ADE56292	Ade56292 Rat Prote
45	143	6.2	864	8 ABO84879	Abo84879 Murine ca

ALIGNMENTS

RESULT 1  
AAB81070  
ID AAB81070 standard; protein; 435 AA.  
XX  
AC AAB81070;  
XX  
DT 25-JUN-2001 (first entry)  
XX  
DE Murine mast cell-specific signal transduction protein.  
XX  
KW Mast cell; signal transduction; mouse; allergic disease.  
XX  
OS Mus musculus.  
XX  
PN JP3146204-B1X  
PD 12-MAR-2001.  
XX  
PF 17-SEP-1999; 99JP-00263778.  
XX  
PR 17-SEP-1999; 99JP-00263778.  
XX  
PA (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.  
XX  
DR WPI; 2001-310022/33.  
DR N-PSDB; AAF86139.  
XX  
PT Mast cell-specific signal-transduction molecule, useful for screening  
therapeutic compounds for treating allergies, is specifically expressed  
by mouse mast cell.  
XX  
PS Claim 1; Page 7-8; 12pp; Japanese.  
XX  
CC This sequence represents a murine mast cell-specific signal transduction  
protein. The invention includes the cDNA and protein sequences of the  
mast cell-specific signal transduction molecule and an expression vector  
containing the polynucleotide sequence. The coding sequence of the signal  
transduction protein can be used for screening therapeutic compounds  
which will be useful for treating allergic diseases  
XX  
SQ Sequence 435 AA;

Query Match 100.0%; Score 2316; DB 4; Length 435;  
Best Local Similarity 100.0%; Pred. No. 1.9e-193;  
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MTSQGNKRTTKKEFGDLRFQVLSLLKNRSPSSAKGRCAVLEPLPDHRRNLAVPGG 60

Db 1 MTSGGNKRTTKEGFDLRFQNVSLKNRSPFSSAKGRCAVLPLEPDRHRLNLAGVPGG 60  
QY 61 EKCSNNNDYEDPBFQLLKAWPSMKILPARPIQSEYADTRYFQDMEAPLPPKASVST 120  
Db 61 EKCSNNNDYEDPBFQLLKAWPSMKILPARPIQSEYADTRYFQDMEAPLPPKASVST 120  
QY 121 EROTDRVMTQLEVDKPTFKDVRQSRQFQKFKYTKINKTLPPLPPRAITLPPKYQPLPPA 180  
Db 121 EROTDRVMTQLEVDKPTFKDVRQSRQFQKFKYTKINKTLPPLPPRAITLPPKYQPLPPA 180  
QY 181 PPESAYFAPKPTFPFVQVQRPQRSAKDFSRVLGAEEESHHTKPESSCPSSNQNTOKS 240  
Db 181 PPESAYFAPKPTFPFVQVQRPQRSAKDFSRVLGAEEESHHTKPESSCPSSNQNTOKS 240  
QY 241 PPAIASSYMPGKHISQIARDHTGSMQHCPCQAORCOAAASHSPRLPYENTNSEKPDPTKPD 300  
Db 241 PPAIASSYMPGKHISQIARDHTGSMQHCPCQAORCOAAASHSPRLPYENTNSEKPDPTKPD 300  
QY 301 EKDVQWNEWYIGEYSRQAVEDVLMKENKDGTFVLVDCSTKSKAEPYVLVYFGNKVYNVK 360  
Db 301 EKDVQWNEWYIGEYSRQAVEDVLMKENKDGTFVLVDCSTKSKAEPYVLVYFGNKVYNVK 360  
QY 361 IRFLESNQOQFALGTGLRGNEFMFDSVEDIIEHYTYFPILLIDGKDKAARRKQCYLTQPLPL 420  
Db 361 IRFLESNQOQFALGTGLRGNEFMFDSVEDIIEHYTYFPILLIDGKDKAARRKQCYLTQPLPL 420  
QY 421 ARLLLTQYSSQALHE 435  
Db 421 ARLLLTQYSSQALHE 435

RESULT 2  
AAE22609  
ID AAE22609 standard; protein; 428 AA.  
XX AC AAE22609;  
XX AC AAE22609;  
DT 26-JUL-2002 (first entry)  
DE Human MIST splice variant protein from clone #7.  
KW Human; mast cell immunoreceptor signal transducer; MIST; immune disorder;  
KW lymphocyte; lymphoma; tumour; thymoma; immune cell disorder; neoplasm;  
KW inflammation; cancer; gene therapy; cytostatic; immunomodulatory.  
XX OS Homo sapiens.  
XX PN WO200226986-A2.  
XX PD 04-APR-2002.  
XX 28-SEP-2001; 2001WO-US030593.  
XX PF 29-SEP-2000; 2000US-0237030P.  
XX PA (BRIM ) BRISTOL-MYERS SQUIBB CO.  
XX PI Perez-Villar JJ, Chang H, Yang W, Wu Y, Whitney GS, Kanner SB;  
XX WPI; 2002-372126/40.  
DR N-PSDB; AAD35801.  
XX PT New isolated mast cell immunoreceptor signal transducer polypeptide,  
PT useful for treating immune disorder involving hyperactivity of B- or T-  
PT lymphocytes, and for inhibiting T-lymphoma, tumor or thymoma.  
XX PS Claim 11; Fig 5; 171pp; English.  
XX The present invention relates to novel mast cell immunoreceptor signal  
transducer (MIST) proteins and polynucleotides encoding such proteins.  
CC MIST sequences of the invention are useful for preventing, treating or  
CC ameliorating a medical condition in mammalian subject. They are useful

CC for treating an immune disorder involving hyperactivity of B- or T-  
CC lymphocytes in a mammal, for inhibiting growth of or inhibiting T-  
CC lymphoma, tumour or thymoma in a mammal. MIST sequences are useful as  
CC targets for therapeutic intervention in immune cell disorders and  
CC inflammatory indicated, for diagnosis and/or screening of disorders or  
CC diseases associated with expression of MIST, for screening for  
CC antagonists or inhibitors of the interaction of MIST with cellular  
CC signalling components. They are used in assays that detect activation or  
CC induction of various B and T-cell-related neoplasms or cancers. Sequences  
CC of the invention are also used in gene therapy. The present sequence is  
CC human MIST splice variant protein from clone #7

SQ Sequence 428 AA;  
Query Match 56.8%; Score 1316.5; DB 5; Length 428;  
Best Local Similarity 62.1%; Pred. No. 4.4e-106;  
Matches 266; Conservative 46; Mismatches 111; Indels 5; Gaps 4;  
QY 1 MTSGGNKRTTKEGFDLRFQNVSLKNRSPFSSAKGRCAVLPLEPDRHRLNLAGVPGG 60  
Db 1 MNRQGNKRTTKEGSNDLKFNFLPKNRSWPRINSATGQYQRMNKPLLDWERNFAAVLDG 60  
QY 61 EKCSNNNDYEDPBFQLLKAWPSMKILPARPIQSEYADTRYFQDMEAPLPPKASVST 120  
Db 61 AKGHSDDDDYDDPELRMEETWQSIKILPARPIKESEYADTHYFKVAMDTPLPLDTRTSISI 120  
QY 121 EROTDRVMTQLEVDKPTFKDVRQSRQFQKFKYTKINKTLPPLPPRAITLPPKYQPLPPA 180  
Db 121 GQPTWNTQ-TRLERVDKPIKSDVRQSNIKGDASVRKNKIPLPPRPLITLPPKYQPLPP- 178  
QY 181 PPESAYFAPKPTFPFVQVQRPQRSAKDFSRVLGAEEESHHTKPESSCPSSNQNTOKS 240  
Db 179 EPESRRPPLSQRHTFPFVQVQRMPSQISLRDLSEVLEAEKVPHNQKPESTHLENQNTQEI 238  
QY 241 PPAIASSYMPGKHISQIARDHTGSMQHCPCQAORCOAAASHSP--RMLPYENTNSEKPDPTK 298  
Db 239 PLAISSTTSNHSVQNRDRHGMQPCSPQRCPPASCSPHENILPYKTSWRPPPKR 298  
QY 299 PDEKDVQWNEWYIGEYSRQAVEDVLMKENKDGTFVLVDCSTKSKAEPYVLVYFGNKVYN 358  
Db 299 SDRKDVQWNEWYIGEYSRQAVEAFMKNENKDGSLVDCSTKSKAEPYVLVYFGNKVYN 358  
QY 359 VKIRFLESNQOQFALGTGLRGNEFMFDSVEDIIEHYTYFPILLIDGKDK-AARRKQCYLTQ 417  
Db 359 VKIRFLESNQOQFALGTGLRGDEKFDSEDIIEHYTKNFPILLIDGKDKTGVRHKQCHLTQ 418  
QY 418 LPLARLL 425  
Db 419 LPLTRHLL 426

RESULT 3  
ADB65190  
ID ADB65190 standard; protein; 428 AA.  
XX AC ADB65190;  
XX AC ADB65190;  
DT 04-DEC-2003 (first entry)  
DE Human protein encoded by clone SPLEN20191020.  
DE DE  
KW Human; pharmaceutical; diagnostic; gene therapy; tissue regeneration;  
KW cell regeneration; membrane protein; signal transduction-related protein;  
KW transcription-related protein; osteoporosis; neurological disease;  
KW cancer; tumour.  
XX OS Homo sapiens.  
XX PN EPI308459-A2.  
XX PD 07-MAY-2003.  
XX 28-MAR-2002; 2002EP-00007401.

```
XX 05-NOV-2001; 2001JP-00379298.
PR 25-JAN-2002; 2002US-00350978.
XX (HELI-) HELIX RES INST.
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX WPI: 2003-450961/43.
DR N-PSDB; ADB63220.
XX
XX New polynucleotides and polypeptides, useful for developing a diagnostic
PT marker or medicines for regulation of their expression and activity, or
PT as targets of gene therapy.
XX
XX Claim 1; Page; 222pp; English.
XX
XX The invention discloses a polynucleotide comprising a sequence selected
CC from 1970 fully defined nucleotide sequences which encode novel
CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
CC or its partial peptide, an antibody binding to the polypeptide or peptide
CC of the polynucleotide, immunologically assaying the polypeptide or
CC peptide of the polynucleotide by contacting the polypeptide or peptide
CC with the antibody of the encoded protein, and observing the binding
CC between the two, a transformant carrying the polynucleotide in an
CC expressible manner and an antisense polynucleotide. The oligonucleotide
CC is useful as a primer for synthesizing the polynucleotide, or as a probe
CC for detecting the polynucleotide. The polynucleotides and encoded
CC proteins are useful as pharmaceutical agents and many disease-related
CC genes may be included in them, for developing a diagnostic marker or
CC medicines for regulation of their expression and activity, or as targets
CC of gene therapy. The genes are involved in tissue and/or cell
CC regeneration. Membrane proteins, signal transduction-related proteins,
CC transcription-related proteins, disease-related proteins and genes
CC encoding them can be used as indicators for diseases (e.g. osteoporosis,
CC neurological diseases, cancer, tumours). The cDNA may be used to regulate
CC the activity or expression of the encoded protein to treat diseases. The
CC sequence presented is a protein of the invention. Note: Some of the
CC sequence data for this patent is not represented in the printed
CC specification, but is based on sequence information supplied by the
CC European Patent Office.
XX
XX Sequence 428 AA;
SQ
Query Match 56.8%; Score 1316.5; DB 7; Length 428;
Best Local Similarity 62.1%; Pred. No. 4.4e-106;
Matches 266; Conservative 46; Mismatches 111; Indels 5; Gaps 4;
QY 1 MTSQGNKRTTKGFGDLRFQNVLLKNRSPWLSAKGRCAVLEPLPDHRRNLGVPGG 60
Db 1 MNRQGNKRTTKGSDNLKFNQFSLPKNRSPRINSATGQYQRMNKPILLDERNFAAVLDG 60
QY 61 EKCSNNNDYDPEPQLKAWPMKILPAPRQESYADRYFQDMMEAPILLPPKASVST 120
Db 61 AKGHSDDYDPELMEETWQSIKILPAPRQESYADRYFQDMMEAPILLPPKASVST 120
QY 121 BRQTRDVRMTQLEVDYPTKDVRSQRFKGYTKINKTLPPLPPRAITIPKKYQPLPPA 180
Db 121 GQPTWNTQ-TRELRVDVPIKSDVRSQIKGASVRKNKIPPLPPRPLITLPKKYQPLPP- 178
QY 181 PPESSAFYAPKPTFPVQVQPRQSAKDFSRVLGAEESHQHPKSSCPSSNQNTQKS 240
Db 179 EPSSRRPLSQRHPTFPEVQMPQSIURDLSEVLEAKVPHNQKRPSTHLENQNTQEI 238
QY 241 PPAIASSYMPGKHSIQARDHTGSMQHCQAQRQAAASHSP--RMLPYENTNSKPDPTK 298
Db 239 PLAISSSSFTTSNHSVQNRDRGGQPCSPQRCQPPASCSPHENILPKYKTSWRPPFKR 298
QY 299 PDEKDVQWQNYIGEYSRQAVEDVLMKENKDGTFVLVDCSTKSKAEYPVLVFGNKVYN 358
Db 299 SDRKDVQWQNYIGEYSRQAVBEAFPMKENKDGSLVDCSTKSKKEYPVLVAFYENKVYN 358
QY 359 VKIRFLESNQOQFALGTGLRGNEPMFDSVEDIIHYTYPPILLIDGKDK-AARRKQCYLTOP 417
Db 359 VKIRFLERNQOQFALGTGLRGDEKFDSDVEDIIHYKNFPIILIDGKDKTGVRHKKOCHLTOP 418
QY 418 LPLARULL 425
Db 419 LPLTRHLL 426
XX
RESULT 4
AAE22608
ID AAE22608 standard; protein; 443 AA.
XX
XX AAE22608;
XX
XX 26-JUL-2002 (first entry)
XX
XX Human MIST protein #1.
XX
XX Human; mast cell immunoreceptor signal transducer; MIST; immune disorder;
KW lymphocyte; lymphoma; tumour; thymoma; immune cell disorder; neoplasm;
KW inflammation; cancer; gene therapy; cytostatic; immunomodulatory.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Modified-site 84
XX Modified-site 111 /label= Tyrosine phosphorylation_binding_site
XX Modified-site 306.311 /label= Tyrosine phosphorylation_binding_site
XX Domain 324.407 /label= SH3_binding_proline-rich_motif
XX Domain /label= SH2_domain
XX
XX WO200226986-A2.
XX
XX 04-APR-2002.
XX
XX 28-SEP-2001; 2001WO-US030593.
XX
XX 29-SEP-2000; 2000US-0237030P.
XX
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
XX Perez-Villar JJ, Chang H, Yang W, Wu Y, Whitney GS, Kanner SB;
XX
XX WPI; 2002-372126/40.
XX
XX N-PSDB; AAD35800.
XX
XX New isolated mast cell immunoreceptor signal transducer polypeptide,
XX useful for treating immune disorder involving hyperactivity of B- or T-
XX lymphocytes, and for inhibiting T-lymphoma, tumor or thymoma.
XX
XX Claim 11; Fig 3; 171pp; English.
XX
XX The present invention relates to novel mast cell immunoreceptor signal
XX transducer (MIST) proteins and polynucleotides encoding such proteins.
XX MIST sequences of the invention are useful for preventing, treating or
XX ameliorating a medical condition in mammalian subject. They are useful
XX for treating an immune disorder involving hyperactivity of B- or T-
XX lymphocytes in a mammal, for inhibiting growth of or inhibiting T-
XX lymphoma, tumour or thymoma in a mammal. MIST sequences are useful as
XX targets for therapeutic intervention in immune cell disorders and
XX inflammatory indications, for diagnosis and/or screening of disorders or
XX diseases associated with expression of MIST, for screening for
XX antagonists or inhibitors of the interaction of MIST with cellular
XX signalling components. They are used in assays that detect activation or
XX induction of various B and T-cell-related neoplasms or cancers. Sequences
XX of the invention are also used in gene therapy. The present sequence is
XX human MIST protein
```

```

XX
SQ Sequence 443 AA;
Query Match 56.8%; Score 1316.5; DB 5; Length 443;
Best Local Similarity 62.1%; Pred. No. 4.6e-106;
Matches 266; Conservative 46; Mismatches 111; Indels 5; Gaps 4;
QY 1 MTSQGNKRTTKGFGDLRFONFVILKNRSPSLSSAKGRCAVLEPLPDHRRNLGAVPGG 60
DB 16 MNQGNKRTTKGSDNDLKFNFLPKNRSWPRINSATGQYQRMKNKPLLDWERNFAAVLDG 75
QY 61 EKNNSNDYEDPFFQLLKAWPMSMKILPARPIQSEYADTRYFQDMMEAPLLLPKASVST 120
DB 76 AKGHSDDDDYDDELRMEETWQSIKILPARPIKSEYADTHYFKVAMDTPLDLTRTSISI 135
QY 121 EROTDRVMTQLEBEVDKPTFKDVRQRFKGYTKINKTLPPLPPRAITLPPKYQPLPPA 180
DB 136 GQPTWNTQ-TRLERVDKPIKSDVRSONIKGDASVRKNKIPLPPRPLITLPPKYQPLPP- 193
QY 181 PPESSAYFAPKPTFEVQGRQRSADKPSRVLGAEESHQTKPESSCPSSNQNTQKS 240
DB 194 EPSSRPPLUSQRHTFEVQRMPSQISLRDLSEVLEAEKVPHNQRPSTHLENQNTQEI 253
QY 241 PPAIASSSYMPGKHSIQARDHTGSMOHCAPQRCQAAASHSP--RMLPYENTNSEKDPDK 298
DB 254 PLAISSSFTTSHSVQNRDRHGMQPCSPQRCQPPASCSPHENILPYKTSWRPFPKR 313
QY 299 PDEKDVQNEWYIGEYSRQAVEDVLMKENKDGTFVLVDCSTKSKAEPYLVVYGNKVYN 358
DB 314 SDRKDVQNEWYIGEYSRQAVEAFMKNKDGSLVRDCSTKSKAEPYLVVYGNKVYN 373
QY 359 VKIRFLESNQOQFALGTCLRGNEFDSVEDIIIEHYTFPILLIDGDK-AARRKQCVLTQ 417
DB 374 VKIRFERNQOQFALGTCLRGDEKFDSEVEDIIIEHYKNFPFILLIDGDKTGVHRKQCHLTQ 433
QY 418 LPLRLALL 425
DB 434 LPLTRHLL 441
RESULT 5
AAB81071
ID AAB81071 standard; protein; 376 AA.
AC AAB81071;
XX
DT 25-JUN-2001 (first entry)
DE Human mast cell-specific immunoreceptor signal transducer.
XX
KW Mast cell; signal transduction; human; allergic disease; MIST;
KW mast cell-specific immunoreceptor signal transducer.
XX
OS Homo sapiens.
XX
PN JP3146204-B1.
XX
PD 12-MAR-2001.
XX
PF 17-SEP-1999; 99JP-00263778.
XX
PR 17-SEP-1999; 99JP-00263778.
XX
PA (KAGA-) KAGAKU GIUTSU SHINKO JIGYODAN.
XX
DR WPI; 2001-310022/33.
XX
DR N-PSDB; AAF86140.
XX
PT Mast cell-specific signal-transduction molecule, useful for screening
PT therapeutic compounds for treating allergies, is specifically expressed
PT by mouse mast cell.
XX
XX Example 1; Page 10-11; 12pp; Japanese.
PS

```

XX This sequence represents human MIST (mast cell-specific immunoreceptor signal transducer). The invention relates to cDNA encoding a murine mast cell-specific signal transduction protein. Included in the invention are cDNA and protein sequences of the mast cell-specific signal transduction molecule and an expression vector containing the polynucleotide sequence. The coding sequence of the signal transduction protein can be used for screening therapeutic compounds which will be useful for treating allergic diseases

XX SQ Sequence 376 AA;

Query Match 48.9%; Score 1132; DB 4; Length 376;  
Best Local Similarity 60.9%; Pred. No. 5e-90;  
Matches 229; Conservative 42; Mismatches 101; Indels 4; Gaps 3;

QY 19 FQNVSLKNRSPSLSSAKGRCAVLEPLPDHRRNLGAVPGGKCNNDYEDPEFOLLK 78  
DB 1 FQNFSLPKNRSWPRINSATGQYQRMKNKPLLDWERNFAAVLDGAKGHSDDDDYDDELRMEE 60

QY 79 AWPMSKILPARPIQSEYADTRYFQDMMEAPLLLPKASVSTERTQDRVMTQLEEVDPK 138  
DB 61 TWQSIKILPARPIKSEYADTHYFKVAMDTPLDLTRTSISIGQPTWNTQ-TRLERVDKP 119

QY 139 TFKDVRSQRFKGYTKINKTLPPLPPRAITLPPKYQPLPPAPPESSAYFAPKPTPEV 198  
DB 120 ISRDVRSQNIKGDASVRKNKIPLPPRPLITLPPKYQPLPP-EPESRRPPLSQRHTPPEV 178

QY 199 QRQPRORSADKPSRVLGAEESHQTKPESSCPSSNQNTQKSPPAIASSSYMPGKHSIQ 258  
DB 179 QGMPSQISLRDLSEVLEAEKVPHNQRPSTHLENQNTQEIPLAISSSFTTSHSVQN 238

QY 259 RDIHTGSMOHCAPQRCQAAASHSP--RMLPYENTNSEKDPDKDEKDVQNEWYIGEYSR 316  
DB 239 RDHRGGMQPCSPQRCQPPASCSPHENILPYKTSWRPFPKRSDRDKDVQNEWYIGEYSR 298

QY 317 QAVEDVLMKENKDGTFVLVDCSTKSKAEPYLVVYGNKVYNKIRFLESNQOQFALGTGL 376  
DB 299 QAVEEAFMKNKDGSLVRDCSTKSKAEPYLVVYGNKVYNKIRFERNQOQFALGTGL 358

QY 377 RGNEMFDSVEDIIIEHY 392  
DB 359 RGDKEKFDSEVEDIIIEHY 374

RESULT 6  
AAE22610  
ID AAE22610 standard; protein; 353 AA.  
AC AAE22610;  
XX  
DT 26-JUL-2002 (first entry)  
DE Human MIST splice variant protein from clone #12.  
XX  
DE Human; mast cell immunoreceptor signal transducer; MIST; immune disorder;  
KW lymphocyte; lymphoma; tumour; thymoma; immune cell disorder; neoplasm;  
KW inflammation; cancer; gene therapy; cytostatic; immunomodulatory.  
XX  
OS Homo sapiens.  
XX  
PN WO200226986-A2.  
XX  
PD 04-APR-2002.  
XX  
PF 28-SEP-2001; 2001WO-US030593.  
XX  
PR 29-SEP-2000; 2000US-0237030P.  
XX  
PA (BRIM ) BRISTOL-MYERS SQUIBB CO.  
XX  
PI Perez-Villar JJ, Chang H, Yang W, Wu Y, Whitney GS, Kanner SB;  
XX

DR WPI; 2002-372126/40.  
 XX N-PSDB; AAD35802.  
 XX New isolated mast cell immunoreceptor signal transducer polypeptide,  
 PT useful for treating immune disorder involving hyperactivity of B- or T-  
 PT lymphocytes, and for inhibiting T-lymphoma, tumor or thymoma.  
 XX Claim 11; Fig 8; 17lpp; English.  
 PS  
 XX The present invention relates to novel mast cell immunoreceptor signal  
 CC transducer (MIST) proteins and polynucleotides encoding such proteins.  
 CC MIST sequences of the invention are useful for preventing, treating or  
 CC ameliorating a medical condition in mammalian subject. They are useful  
 CC for treating an immune disorder involving hyperactivity of B- or T-  
 CC lymphocytes in a mammal, for inhibiting growth of or inhibiting T-  
 CC lymphoma, tumour or thymoma in a mammal. MIST sequences are useful as  
 CC targets for therapeutic intervention in immune cell disorders and  
 CC inflammatory indications, for diagnosis and/or screening of disorders or  
 CC diseases associated with expression of MIST, for screening for  
 CC antagonists or inhibitors of the interaction of MIST with cellular  
 CC signalling components. They are used in assays that detect activation or  
 CC induction of various B and T-cell-related neoplasms or cancers. Sequences  
 CC of the invention are also used in gene therapy. The present sequence is  
 CC human MIST splice variant protein from clone #12  
 XX  
 SQ Sequence 353 AA;  
 Query Match 47.9%; Score 1109.5; DB 5; Length 353;  
 Best Local Similarity 64.4%; Pred. No. 4.2e-88;  
 Matches 226; Conservative 31; Mismatches 89; Indels 5; Gaps 4;  
 QY 78 KAWPSMKILPARIQSEYADTYFQDMMEAPILLPPKASVSTERQTRDVRMTQLEVDK 137  
 Db 3 ETQSIKILPARIQSEYADTYFQDMMEAPILLPPKASVSTERQTRDVRMTQLEVDK 61  
 QY 138 PTFKDVRSORFKYTKINKTLPPLPPPAITLPPKYPQPLPAPPRESSAYFAPKTFPE 197  
 Db 62 PISKDVRSQNIKGDASVRKNKILFLPPPRPLITLPPKYPQPLP-EPSSRRPLSQRTFPE 120  
 QY 198 VQGPQRORSADKFSRLVGAEEESHQTPKSPSSNQNTQKSPALAIASSYMPGKHISIQ 257  
 Db 121 VQMPQSISRLDSEVLEAKVPHNQKPESTHLENQNTQETPLAISSSFTTSHSVQ 180  
 QY 258 ARDHTGSMQHCAPQRCQAASHSP--RMLPYNTNSSEKPDTPKDEKVMQNEWYIGEYS 315  
 Db 181 NRDRGGMQPCSPQRCQPPASCSPHENILPKYTSWRPFPKRSRDKDVQHNEWYIGEYS 240  
 QY 316 QRAVEDVLMKENKDGFLVDFDCSTKSAEPYLVVFGNKVYKVRFLFESNQPFALGTG 375  
 Db 241 QRAVEEAFMKENKDGSLVDFDCSTKSEEPYLVAVFYENKYNKIRFLERNQPFALGTG 300  
 QY 376 LRGNEMFDSVEDIEHYTFYPIILLIDGDK-AARRKQCYLTQPLPLARLLL 425  
 Db 301 LRGDEKFDSDVEDIEHYKNFPIILLIDGDKDTGVHRKQCHLTQPLPLTRHLL 351  
 RESULT 7  
 ADP65191  
 ID ADP65191 standard; protein; 533 AA.  
 XX  
 AC ADP65191;  
 XX  
 DT 12-AUG-2004 (first entry)  
 XX  
 DE Human lymphocyte cytosolic protein 2, SH2 domain-containing leukocyte.  
 XX  
 KW autoimmune disease; arthritis; gene expression analysis;  
 KW rheumatoid arthritis; collagen-induced; immunosuppressive; anti-rheumatic;  
 KW antiarthritic; osteopathic; antigout; antiinflammatory; dermatological;  
 KW immunomodulatory; lupus; ankylosing spondylitis; fibrositis;  
 KW fibromyalgia; osteoarthritis; gout; juvenile rheumatoid arthritis;  
 KW immune; human.  
 XX

OS Homo sapiens.  
 XX WO2003072827-A1.  
 XX 04-SBP-2003.  
 XX 31-OCT-2002; 2002WO-US035433.  
 XX 31-OCT-2001; 2001US-0336220P.  
 XX (CHIL-) CHILDREN'S HOSPITAL MEDICAL CENT.  
 XX Hirsch R, Thorton SL;  
 XX WPI; 2003-712740/67.  
 XX GENBANK; NP\_005556.  
 XX Diagnosing and analyzing autoimmune disease using gene expression  
 PT profiles and microarray technology, useful for diagnosing and treating  
 PT rheumatoid arthritis, lupus, fibrositis, osteoarthritis, fibromyalgia and  
 XX gout.  
 PS Disclosure; Page; 56pp; English.  
 XX The invention relates to a novel method for diagnosing and analysing  
 CC autoimmune disease or arthritides. The method comprises obtaining a  
 CC patient sample containing mRNA, analysing gene expression using the mRNA  
 CC that results in a gene expression signature of the mRNA, and using that  
 CC gene expression signature to diagnose or analyse the autoimmune disease  
 CC or arthritides in the patient, where gene expression of at least 60% of  
 CC the genes correlates with that of the gene signature. The invention  
 CC further comprises: a treatment of rheumatoid arthritis; identification of  
 CC genes for targeting in the treatment of rheumatoid arthritis in a mammal  
 CC other than a mouse; diagnosis of rheumatoid arthritis in a mammal; an  
 CC array or gene chip, specific for rheumatoid arthritis; diagnosis or  
 CC analyses of autoimmune disease or rheumatoid arthritis; screening the  
 CC efficacy of a candidate drug in vitro for the treatment of collagen-  
 CC induced arthritis; and reducing the symptoms associated with collagen-  
 CC induced arthritis. The compositions of the invention have the following  
 CC activities: immunosuppressive, anti-rheumatic, antiarthritic, osteopathic,  
 CC antigout, antiinflammatory, dermatological, and immunomodulatory. The  
 CC methods and compositions of the present invention are useful for  
 CC diagnosing and treating autoimmune disease or arthritides, such as  
 CC rheumatoid arthritis, lupus, ankylosing spondylitis, fibrositis,  
 CC fibromyalgia, osteoarthritis, gout, juvenile rheumatoid arthritis, and an  
 CC immune disease caused by an infectious agent. This sequence represents a  
 CC protein sequence relating to the genes used in the analysis and treatment  
 CC of autoimmune diseases or arthritides. Note: This sequence is not shown  
 CC in the specification. It has been supplied in an electronic format from  
 XX WIPO.  
 SQ Sequence 533 AA;  
 Query Match 17.2%; Score 398; DB 7; Length 533;  
 Best Local Similarity 30.2%; Pred. No. 1.2e-25;  
 Matches 124; Conservative 52; Mismatches 146; Indels 88; Gaps 14;  
 QY 68 DYEDPEFQLLKAWPSMKILPARIQSE--YADTRYFQDMMEAP-----LLLPKKA 116  
 Db 144 DYEPFNSDEALQN-SILPAKFPNNSNMYIDRPPSGKTPQQPPVPPQRPMAALPPPPA 202  
 QY 117 -----SVSTERQTRDVRMTQLEVD-----KPT 139  
 Db 203 GRNHSPLPPQTNEHPPSRNHNKTAKLPAPSIDRSTKPLPLDRSLAPDFEPFTLGKKPP 262  
 QY 140 FKD---VRSORFGKYTKINKTLPPLPPPAITLPPKYPQPLPAPPRESSAYFAPKTFPE 196  
 Db 263 FSDKPSIPAGRSUHEHLPKIQKPLPP-----TTERHSSSLP-----GKKPVP 308  
 QY 197 EVQRGPRQRSADKFSRLVGAEEESHQTKPE-----SSCFSSNQNTQKSPALAISSSY 250  
 Db 309 KHGWGPDRE-----NDEDDVHQRPPLQPALLPMSSNTFFSRSTKPSPMNPLPSSHM 360

QY 251 PGKHSIQARDHTGSMQHCQAQRAASHSP-----RMLPYENTNSEKPDPTKDEKDVW 305  
 Db 361 PGAFS-ESNSFFQSASLPYFSGPSNRPPPIRAEGRNFFPLPNKPRP-PSPAEENS 418  
 QY 306 QNEWYIGEYGRQAVEDVLMKENKDGTFVLVDCSTKSKAEYPVLVVFYGNKVNKIRFLE 365  
 Db 419 NEEWYYSYITRPEAEALRKINQDGTFLVRDSSKTTTNPYVLVLYKDKVNIQIRYQK 478  
 QY 366 SNOQFALGTGLRGNEMFDSVEDIEHYTFPIILLIDGKDKAARKQCYLT 415  
 Db 479 ESQVYLLGTGLRGKEDFLSVSDIIDYFRKMPILLIDGKNRGS-RYQCTLT 527

RESULT 8  
 ADQ19792  
 ID ADQ19792 standard; protein; 533 AA.  
 XX AC ADQ19792;  
 XX DT 26-AUG-2004 (first entry)  
 XX DE Human soft tissue sarcoma-upregulated protein - SEQ ID 2611.  
 XX KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.  
 XX OS Homo sapiens.  
 XX PN WO2004048938-A2.  
 XX PD 10-JUN-2004.  
 XX PF 26-NOV-2003; 2003WO-US038193.  
 XX PR 26-NOV-2002; 2002US-0429739P.  
 XX PA (PROT-) PROTEIN DESIGN LABS INC.  
 XX PI Aziz N, Ginsburg WM, Zlotnik A;  
 XX WPI; 2004-441208/41.  
 XX PT Early detection of soft tissue sarcoma comprises determining expression  
 PT of a gene in a first soft tissue sample and a normal soft tissue sample  
 PT and comparing the gene expression, also useful in treating soft tissue  
 PT sarcoma.  
 XX PS Example 2; SEQ ID NO 2611; 210pp; English.  
 CC The invention relates to a novel method for detecting soft tissue sarcoma  
 CC which comprises obtaining a first soft tissue sample from an individual  
 CC and a normal soft tissue sample from the same or different individual,  
 CC determining the expression of a gene in both samples and comparing the  
 CC expression of the gene in both soft tissue samples, where a higher level  
 CC of protein expression in the first soft tissue sample indicates the  
 CC presence of soft tissue sarcoma. The method of the invention has  
 CC cytostatic applications and may be useful for detecting soft tissue  
 CC sarcoma, possibly via gene therapy or vaccine production. The nucleic  
 CC acid sequences may be useful in diagnostic and screening applications.  
 CC The current sequence is that of a human soft tissue sarcoma-upregulated  
 CC protein of the invention. The current sequence is not shown within the  
 CC specification per se but was submitted in CD format by the inventor.  
 XX Sequence 533 AA;  
 SQ Query Match 17.2%; Score 398; DB 8; Length 533;  
 Best Local Similarity 30.2%; Pred. No. 1.2e-25;  
 Matches 124; Conservative 52; Mismatches 146; Indels 88; Gaps 14;  
 QY 68 DYEDPEQLLKAWPSMKILPARIQSE--YATRYFQDMMEAP-----LLLPKKA 116  
 Db 144 DYEPPPSNDEALQN-SILPAKPFPSNSMIDRPFSKGTPOQPPVFPQRPMAALPPPPA 202  
 QY 117 -----SVSTERQTRDVRMTQLEVD-----KPT 139

Db 203 GRNHSPLPPQTTHBPSRSRNHKTAKLPAPSIDRSTKPLDRLSLAPDFREPFTLGKKPP 262  
 QY 140 FKD---VRSORFXGFKYTKINTPLPPRPAITLPKKYQPLPPAPPESAYFAPKFTFP 196  
 Db 263 FSDKPSIPAGRSIGEHLPKIKPPLP-----TTERHERSSPLP-----GKKDPVP 308  
 QY 197 EVQGRPRQRAKDFSRVLGAEBESHQTKPE-----SSCFSSNQNTQKSPPAIASSYM 250  
 Db 309 KHGWSGDRRE-----NDEDDVHORPLPQPALLPMSSNTFPSSRTKPSPMNPLPSSHM 360  
 QY 251 PGKHSIQARDHTGSMQHCQAQRAASHSP-----RMLPYENTNSEKPDPTKDEKDVW 305  
 Db 361 PGAFS-ESNSFFQSASLPYFSGPSNRPPPIRAEGRNFFPLPNKPRP-PSPAEENS 418  
 QY 306 QNEWYIGEYGRQAVEDVLMKENKDGTFVLVDCSTKSKAEYPVLVVFYGNKVNKIRFLE 365  
 Db 419 NEEWYYSYITRPEAEALRKINQDGTFLVRDSSKTTTNPYVLVLYKDKVNIQIRYQK 478  
 QY 366 SNOQFALGTGLRGNEMFDSVEDIEHYTFPIILLIDGKDKAARKQCYLT 415  
 Db 479 ESQVYLLGTGLRGKEDFLSVSDIIDYFRKMPILLIDGKNRGS-RYQCTLT 527

RESULT 9  
 ADP23847  
 ID ADP23847 standard; protein; 533 AA.  
 XX AC ADP23847;  
 XX DT 18-NOV-2004 (first entry)  
 XX DE PRO polypeptide SEQ ID NO:1025.  
 XX KW PRO; antiinflammatory; antiarthritic; antirheumatic; immunosuppressive;  
 KW osteopathic; antidiabetic; dermatological; antipsoriatic; antiallergic;  
 KW antiasthmatic; hepatotropic; respiratory; gene therapy; immune system.  
 XX OS Unidentified.  
 XX PN WO2004041170-A2.  
 XX PD 21-MAY-2004.  
 XX PF 30-OCT-2003; 2003WO-US034312.  
 XX PR 01-NOV-2002; 2002US-0423394P.  
 XX (GETH ) GENENTECH INC.  
 XX PA Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;  
 PI Wu TD;  
 XX WPI; 2004-419628/39.  
 XX N-PSDB; ADP23846.  
 XX New PRO polypeptides and polynucleotides, useful for treating e.g.  
 PT erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated  
 PT renal disease, or demyelinating diseases of the central or peripheral  
 PT nervous system.  
 XX Claim 7; SEQ ID NO 1025; 2940pp; English.  
 CC The invention relates to a novel isolated nucleic acid and the PRO  
 CC polypeptide encoded by it. A protein of the invention has  
 CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,  
 CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,  
 CC antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide  
 CC of the invention may have a use in gene therapy. The PRO polypeptide, its  
 CC agonist, antagonist, or antibody that specifically binds to the  
 CC polypeptide is useful for treating an immune related disorder such as  
 CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,  
 CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an

```

CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
CC disease, a demyelinating disease of the central or peripheral nervous
CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,
CC a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary
CC disease, infectious or autoimmune chronic active hepatitis, primary
CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
CC disease, an autoimmune or immune-mediated skin disease, a bullous skin
CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic
CC disease, asthma, allergic rhinitis, atopic dermatitis, food
CC hypersensitivity, urticaria, an immunologic disease of the lung,
CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
CC pneumonitis, a transplantation associated disease, graft rejection or
CC graft-versus-host disease. The present sequence represents a PRO protein
CC of the invention.
XX
XX
SQ Sequence 533 AA;
Query Match 17.2%; Score 398; DB 8; Length 533;
Best Local Similarity 30.2%; Pred. No. 1.2e-25;
Matches 124; Conservative 52; Mismatches 146; Indels 88; Gaps 14;
Qy 68 DYEDPEFQLLKAWPSMKILPARPIQSE--YADTRYFQDMWEAP-----LLLPPKA 116
Db 144 DYEPPPSNDEALQN-SILPAKPPFNNSMYIDRPPSGKTPQQPPVPQRPMAALPPPPA 202
Qy 117 -----SVSTERQTRVDMTOLBEVD-----KPT 139
Db 203 GRNHSPLPPQTNHEEPPSRNHNHTAKLPAPSIDRSTKPLDRLSLAPFDREPFTLGKKPP 262
Qy 140 FKD---VRSORFKGFKYTKINKTLPPLPPRPAITLKKYQLPLPAPPBESSAYFAPKPTFP 196
Db 263 FSDKPSIPAGSLGELHLPKIQKPLPP-----TTERHERSSPLP-----GKKPPVP 308
Qy 197 EVQGGPQRSAKDSFVRLGAEESHHQTKPE-----SSCPSSNQNTQKSPPAIASSYM 250
Db 309 KHGSGPDRRE-----NDEDDVHQRPLOPALLPMSSNTTFFSRSTKPSPMNPLPSHHM 360
Qy 251 PGKHSIQARDHTGSMQHCPAQRQAAASHSP-----RMLPVNTNSKPKPTKDEKDVW 305
Db 361 PGAFS-ESNSFFQSASLPPYFSGPSNRPRPIRAEGNFPLPLNPKPRP-PSPAEEENSL 418
Qy 306 QNEWYIGEYGRQAVEDVLMKENKDGTLVRDCSTKSAEPYVLVVFYGNKYVNVKIRFLE 365
Db 419 NEWYVSVYITPEAEALRKINQDGTFLVRDSSKKTTPNPLYMLVKDKYVNIQIRYQK 478
Qy 366 SNOQFALGTGLRGNMFDSVEDIEHTYTFPILLIDGKDKAARKQCVLT 415
Db 479 ESQVYLLGTGLRGKEDFLSVSDIIDYFRKMPLLLLIDGKNRGS-RYQCTLT 527

RESULT 10
ADP23845
AC ADP23845 standard; protein; 533 AA.
XX
XX ADP23845;
XX
XX 18-NOV-2004 (first entry)
XX
XX PRO polypeptide SEQ ID NO:1023.
XX
XX PRO; antiinflammatory; antiarthritic; antirheumatic; immunosuppressive;
XX osteopathic; antidiabetic; dermatological; antipsoriatic; antiallergic;
XX antiasthmatic; hepatotropic; respiratory; gene therapy; immune system.
XX
XX Unidentified.
XX
XX WO2004041170-A2.
XX
XX 21-MAY-2004.
XX
XX

```

```

PF 30-OCT-2003; 2003WO-US034312.
XX
XX 01-NOV-2002; 2002US-0423394P.
XX
XX (GETH ) GENENTECH INC.
XX
XX Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WJ;
XX Wu TD;
XX
XX WPI; 2004-419628/39.
XX
XX N-PSDB; ADP23844.
XX
XX New PRO polypeptides and polynucleotides, useful for treating e.g.
XX erythematosus, rheumatoid arthritis, diabetes mellitus, immune-mediated
XX renal disease, or demyelinating diseases of the central or peripheral
XX nervous system.
XX
XX Claim 7; SEQ ID NO 1023; 2940pp; English.
XX
XX The invention relates to a novel isolated nucleic acid and the PRO
XX polypeptide encoded by it. A protein of the invention has
XX antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,
XX osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,
XX antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide
XX of the invention may have a use in gene therapy. The PRO polypeptide, its
XX agonist, antagonist, or antibody that specifically binds to the
XX polypeptide is useful for treating an immune related disorder such as
XX systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,
XX juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
XX idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
XX vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
XX thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
XX disease, a demyelinating disease of the central or peripheral nervous
XX system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,
XX a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary
XX disease, infectious or autoimmune chronic active hepatitis, primary
XX biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
XX inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
XX disease, an autoimmune or immune-mediated skin disease, a bullous skin
XX disease, erythema multiforme, contact dermatitis, psoriasis, an allergic
XX disease, asthma, allergic rhinitis, atopic dermatitis, food
XX hypersensitivity, urticaria, an immunologic disease of the lung,
XX eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
XX pneumonitis, a transplantation associated disease, graft rejection or
XX graft-versus-host disease. The present sequence represents a PRO protein
XX of the invention.
XX
XX
XX Sequence 533 AA;
Query Match 17.2%; Score 398; DB 8; Length 533;
Best Local Similarity 30.2%; Pred. No. 1.2e-25;
Matches 124; Conservative 52; Mismatches 146; Indels 88; Gaps 14;
Qy 68 DYEDPEFQLLKAWPSMKILPARPIQSE--YADTRYFQDMWEAP-----LLLPPKA 116
Db 144 DYEPPPSNDEALQN-SILPAKPPFNNSMYIDRPPSGKTPQQPPVPQRPMAALPPPPA 202
Qy 117 -----SVSTERQTRVDMTOLBEVD-----KPT 139
Db 203 GRNHSPLPPQTNHEEPPSRNHNHTAKLPAPSIDRSTKPLDRLSLAPFDREPFTLGKKPP 262
Qy 140 FKD---VRSORFKGFKYTKINKTLPPLPPRPAITLKKYQLPLPAPPBESSAYFAPKPTFP 196
Db 263 FSDKPSIPAGSLGELHLPKIQKPLPP-----TTERHERSSPLP-----GKKPPVP 308
Qy 197 EVQGGPQRSAKDSFVRLGAEESHHQTKPE-----SSCPSSNQNTQKSPPAIASSYM 250
Db 309 KHGSGPDRRE-----NDEDDVHQRPLOPALLPMSSNTTFFSRSTKPSPMNPLPSHHM 360
Qy 251 PGKHSIQARDHTGSMQHCPAQRQAAASHSP-----RMLPVNTNSKPKPTKDEKDVW 305
Db 361 PGAFS-ESNSFFQSASLPPYFSGPSNRPRPIRAEGNFPLPLNPKPRP-PSPAEEENSL 418
Qy 306 QNEWYIGEYGRQAVEDVLMKENKDGTLVRDCSTKSAEPYVLVVFYGNKYVNVKIRFLE 365
Db 419 NEWYVSVYITPEAEALRKINQDGTFLVRDSSKKTTPNPLYMLVKDKYVNIQIRYQK 478
Qy 366 SNOQFALGTGLRGNMFDSVEDIEHTYTFPILLIDGKDKAARKQCVLT 415
Db 479 ESQVYLLGTGLRGKEDFLSVSDIIDYFRKMPLLLLIDGKNRGS-RYQCTLT 527

```

QY 306 QNEWYIGEYSRQAVEDVLMKENKDGTFELVRDCSTKSKAEYPVLVVFYGNKVYNKIRFLE 365  
 Db 419 NEWYVSYITRPEAAALRKINQDGTFLVRDSSKTTTNPYVLVLYKDKVNIQIRYQK 478  
 QY 366 SNOQFALGTGLRGNEFDSVEDIEHYTFYPILLIDGKDKAARRKQCYLT 415  
 Db 479 ESQVYLLGTGLRGKEDFLSVSDIIDYFRKMPLLLLIDGKNRGS-RYQCTLT 527

RESULT 11  
 ADT07538 standard; protein; 533 AA.  
 XX ADT07538;  
 XX  
 DT 13-JAN-2005 (first entry)  
 DE Human colon-specific polypeptide (CSP) #55.  
 XX Human; colon-specific polypeptide; CSP; metastasis; colon cancer;  
 KW cytostatic.  
 KW  
 XX Homo sapiens.  
 XX  
 PN W02004089301-A2.  
 XX  
 XX 21-OCT-2004.  
 XX  
 XX 02-APR-2004; 2004WO-US010531.  
 PF  
 XX  
 PR 02-APR-2003; 2003US-0460711P.  
 XX  
 XX (DIAD-) DIADEXUS INC.  
 XX  
 XX Macina RA, Turner LR, Sun Y;  
 XX  
 XX WPI; 2004-748572/73.  
 DR N-PSDB; ADT07466.  
 XX  
 XX New mammalian or human nucleic acid molecule or polypeptide, useful in  
 PT preparing a composition for diagnosing or treating colon cancer.  
 PT  
 XX Claim 12; SEQ ID NO 125; 596pp; English.  
 XX  
 XX The invention relates to the isolation of human colon specific nucleic  
 CC acids (CSNAs) and the encoding colon-specific polypeptides (CSPs). Also  
 CC disclosed is a method for determining the presence of a CSNA in a sample,  
 CC a method for diagnosing or monitoring the presence and metastases of  
 CC colon cancer in a patient, a method of treating a patient with colon  
 CC cancer, a vaccine consisting of the isolated nucleic acid molecule or  
 CC polypeptide, and a kit for detecting a risk of cancer or presence of  
 CC cancer in a patient. The nucleic acid molecule is useful in preparing a  
 CC composition for diagnosing or treating colon cancer. The present sequence  
 CC represents a CSP of the invention.  
 XX  
 XX Sequence 533 AA;  
 SQ

Query Match 17.2%; Score 398; DB 8; Length 533;  
 Best Local Similarity 30.2%; Pred. No. 1.2e-25;  
 Matches 124; Conservative 52; Mismatches 146; Indels 88; Gaps 14;  
 QY 68 DYEDPEFOLLKAWPSMKILPARIQSE--YADTRYFQDMNEAP-----LLLPKKA 116  
 Db 144 DYEPPPPNDSEALQN-SILPAKPPFNNSMYIDRPPSGKTPQQPPVPPQRPMAALPPPPA 202  
 QY 117 -----SVSTEROTRVRMTQLEVD-----KPT 139  
 Db 203 GRNHSPLPPPTQNHSPSRNHNKTAKLPAISIDRSTKPLDRSLAPFDEPFTLGKPP 262  
 QY 140 FKD---VRSORFGFKYTKINKTLPPLPPRPAILTPKKYQPLPPAPPEESSAYFAPKPTFP 196  
 Db 263 FSDKPSIPAGSLGEHLPKIQKPLPP-----TTERHERSSPLP-----GKKPPVP 308

QY 197 EVQGRQRSAKDFSRVLGAEEESHQTKPE-----SSCPSNNQNTQKSPPAATASSYM 250  
 Db 309 KHGWGDPDRR-----NDEDDVHORPLPQALLPMSSNTFFSRSTKPSMNPPLPSHM 360  
 QY 251 PGKHSIQARDHTGSMOHCFAQRCAAAASHSP-----RMLPYENTNSENKPTKPDKDVW 305  
 Db 361 PGAFS-ESNSSFPQSASLPPYFSGPSNRPPIRAEGRNFFPLPLPNKPRP-PSPAEEENSL 418  
 QY 306 QNEWYIGEYSRQAVEDVLMKENKDGTFELVRDCSTKSKAEYPVLVVFYGNKVYNKIRFLE 365  
 Db 419 NEWYVSYITRPEAAALRKINQDGTFLVRDSSKTTTNPYVLVLYKDKVNIQIRYQK 478  
 QY 366 SNOQFALGTGLRGNEFDSVEDIEHYTFYPILLIDGKDKAARRKQCYLT 415  
 Db 479 ESQVYLLGTGLRGKEDFLSVSDIIDYFRKMPLLLLIDGKNRGS-RYQCTLT 527

RESULT 12  
 ADP45459  
 ID ADP45459 standard; protein; 224 AA.  
 XX  
 AC ADP45459;  
 XX  
 DT 09-SEP-2004 (first entry)  
 XX  
 DE Human colon cancer cell-associated protein SeqID182.  
 XX  
 KW neoplastic colon cell; cytostatic; vaccine; gene therapy; colon cancer;  
 KW immune response; metastasis; human.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W02004050900-A2.  
 XX  
 XX 17-JUN-2004.  
 XX  
 XX 04-DEC-2003; 2003WO-US040131.  
 PF  
 XX  
 PR 04-DEC-2002; 2002US-0431143P.  
 PR 04-DEC-2002; 2002US-0431206P.  
 XX  
 XX (DIAD-) DIADEXUS INC.  
 XX  
 XX Macina RA, Turner LR, Sun Y, Rodriguez M;  
 XX  
 XX WPI; 2004-480629/45.  
 DR N-PSDB; ADP45333.  
 XX  
 XX Nucleic acid molecules isolated from normal and neoplastic colon cells,  
 PT useful for treating colon cancer and for identifying, diagnosing,  
 PT monitoring, staging, and imaging colon cancer and/or non-cancerous  
 PT disease states in colon.  
 XX  
 XX Claim 12; SEQ ID NO 182; 564pp; English.  
 XX  
 XX This invention relates to novel nucleic acid molecules isolated from  
 CC normal and neoplastic colon cells and the proteins encoded by them. The  
 CC invention may be useful for the production of compounds with a cytostatic  
 CC activity or for the development of a vaccine or gene therapy. The  
 CC invention may be useful for developing a treatment for a patient with  
 CC colon cancer, where the administration induces an immune response against  
 CC the colon cancer cell expressing the novel nucleic acids or proteins. The  
 CC invention may also be useful for diagnosing or monitoring the presence  
 CC and metastases of colon cancer in a patient. The present sequence is that  
 CC of a colon cancer cell-associated protein of the invention.  
 XX  
 XX Sequence 224 AA;  
 SQ

Query Match 14.6%; Score 337; DB 8; Length 224;  
 Best Local Similarity 40.0%; Pred. No. 7.7e-21;  
 Matches 78; Conservative 33; Mismatches 76; Indels 8; Gaps 4;  
 QY 226 PESSCPSNNQNTQKSPPAATASSYMPGKHSIQARDHTGSMQHCFAQRCAAAASHSP---- 281





[illegible]

Qy	282	-RMLPYENTNGEKPDPTKPDKDVQWQNEWYIGEYSRQAVEDVLMKENKDGTFLVRDCSTK	340
Db	92	GRNFPLPLPNKPRP-PSPAEEENSLENEWYVSYITRPEAAEARLKINQDGTFLVRDSSKK	150
Qy	341	SKAPPYLVAVFYGNKVVNKKTRFLESNOQFALGTGLRGNEMFDSVEDIIIEHYTYPIILLI	400
Db	151	TTTTTPYVLVUJYKDVYVNIQIRYKESQYLLGTGLRGKEDFLSVSDIIDYFRKMPLLLI	210
Qy	401	DGDKKAARRKQCYLT 415	
Db	211	DGKNRGS-RYQCTLT 224	
RESULT	15		
AA	B93406		
ID	AA	B93406 standard; protein; 456 AA.	
XX	XX	AA	B93406;
XX	XX	XX	XX
DT	26-JUN-2001	(first entry)	
XX	XX	Human protein sequence SEQ ID NO:12602.	
DE	DE	Human; primer; detection; diagnosis; antisense therapy; gene therapy.	
KW	KW	Homo sapiens.	
OS	OS	EP1074617-A2.	
PN	PN	07-FEB-2001.	
PD	PD	28-JUL-2000; 2000EP-00116126.	
PF	PF	29-JUL-1999; 99JP-00248036.	
XX	XX	27-AUG-1999; 99JP-00300253.	
PR	PR	11-JAN-2000; 2000JP-00118776.	
PR	PR	02-MAY-2000; 2000JP-00183767.	
PR	PR	09-JUN-2000; 2000JP-00241899.	
XX	XX	(HELIX-) HELIX RES INST.	
PA	PA	Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;	
XX	XX	Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;	
PI	PI	WPI: 2001-318749/34	
XX	XX		
DR	DR		

AA Primer sets for synthesizing polynucleotides, particularly the 5602 full-  
 AT length cDNAs defined in the specification, and for the detection and/or  
 PT diagnosis of the abnormality of the proteins encoded by the full-length  
 PT cDNAs.  
 XX  
 PS Claim 8; SEQ ID NO 12602; 2537pp + Sequence Listing; English.  
 XX  
 CC The present invention describes primer sets for synthesising 5602 full-  
 CC length cDNAs defined in the specification. Where a primer set comprises:  
 CC<sup>1</sup> (a) an oligo-dT primer and an oligonucleotide complementary to the  
 CC complementary strand of a polynucleotide which comprises one of the 5602  
 CC nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the  
 CC specification. The primer sets can be used in antisense therapy and in  
 CC gene therapy. The primers are useful for synthesising polynucleotides, and  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB32446 to AAB95893

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 22, 2006, 01:32:14 ; Search time 41 Seconds  
(without alignments)  
1020.836 Million cell updates/sec

Title: US-10-717-619-2  
Perfect score: 2316  
Sequence: 1 MTSQGNKRTTKGFGDLRFQ.....QLPLARLLLTQYSSQALHE 435

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 80.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	412.5	17.8	533	2 B56110	tyrosine phosphor
2	398	17.2	533	2 A56110	tyrosine phosphor
3	154.5	6.7	663	1 TVMVPR	protein-tyrosine k
4	146	6.3	728	2 H59435	phosphoinositide-3
5	142.5	6.2	723	2 B38749	3-phosphatidylinos
6	138.5	6.0	443	2 T27877	hypothetical prote
7	134.5	5.8	480	2 JC7552	Shb-like adapter p
8	134	5.8	1599	2 T15854	hypothetical prote
9	133	5.7	675	2 S60612	protein-tyrosine k
10	132.5	5.7	592	1 LLBY	actin-binding prot
11	132	5.7	1603	2 S17983	gene posterior sex
12	130.5	5.6	724	2 A38749	3-phosphatidylinos
13	130	5.6	559	2 I49444	SH3 binding protei
14	130	5.6	724	2 A38747	phosphatidylinosit
15	129	5.6	751	2 S68957	adhesive plaque pr
16	128.5	5.5	3421	1 W2B866	367K tegument prot
17	128	5.5	706	2 A45990	junctional sarcopt
18	128	5.5	724	2 A38748	3-phosphatidylinos
19	127	5.5	1385	2 T21706	hypothetical prote
20	125.5	5.4	1029	2 T30351	mucin-like protein
21	125	5.4	461	2 A57463	p55PIK - mouse
22	125	5.4	962	2 S58107	hypothetical WD-40
23	124.5	5.4	473	2 S25776	transforming prote
24	124.5	5.4	585	2 A46209	protein-tyrosine-p
25	124.5	5.4	593	1 JN0805	protein-tyrosine-p
26	124.5	5.4	597	1 A53593	protein-tyrosine-p
27	123.5	5.3	841	2 A43254	protein-tyrosine-p
28	123.5	5.3	1317	2 T03748	apoptosis associat
29	122.5	5.3	596	2 I38228	Shb protein - huma

polyphenolic adhes  
hypothetical prote  
hypothetical prote  
1-phosphatidylinos  
tensin - chicken ( ENL (translocation  
gag-abl polyprotei  
protein-tyrosine k  
protein-tyrosine-p  
ribonuclease E VC2  
tensin - chicken  
tensin, cardiac mu  
hypothetical prote  
nascent polypeptid  
Wiskott-Aldrich sy  
probable protein k

ALIGNMENTS

RESULT 1

B56110  
tyrosine phosphoprotein SLP-76 - mouse  
C/Species: Mus musculus (house mouse)  
C/Date: 28-Apr-1995 #sequence\_revision 28-Apr-1995 #text\_change 09-Jul-2004  
C/Accession: B56110  
R/Jackman, J.K.; Mott, D.G.; Sun, Q.; Tanemoto, M.; Turck, C.W.; Peltz, G.A.; Koretzky, J. Biol. Chem. 270, 7029-7032, 1995  
A/Title: Molecular cloning of SLP-76, a 76-kDa tyrosine phosphoprotein associated with G G  
A/Reference number: A56110; MUID:95221345; PMID:7706237  
A/Accession: B56110  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-533 <JAC>  
A/Cross-references: UNIPROT:Q60787; UNIPARC:UPI00000284C7; GB:U20159; NID:g806767; PIDN:  
C/Keywords: phosphoprotein  
F:422-520/Domain: SH2 homology <SH2>

Query Match	17.8%	Score	412.5	DB 2	Length	533
Best Local Similarity	29.7%	Pred. No.	2.5e-20			
Matches	148	Conservative	51	Mismatches	175	Indels 125; Gaps 20;
Qy	14	FGDLRFQNVSL	-----XNRSWPSLSAKRCRAVLEPLDPRHRLAG-----	56		
Db	57	FPKLRMPLLSKLSQDINKNEERSIFTRKQIPRFLEETESHEEDDCGWSSEFDDYESPN	116			
Qy	57	--VPGGEKCSNNYEDP--EFQLL-----KAWPS-----MKTLPARPIQ--ESEY	96			
Db	117	DDPDGE---DDGDYSPNEEQALVDDAADYEPFPPSNNEALQSSILPPNSFHTNSMY	173			
Qy	97	ADTRYFQDMMEAPLLLP--PKASV-----	118			
Db	174	IDRPPTGKVSQQPPVPLRPKLPALPLPTGHNHSPSPHPNHEEPSRSGNNKTAKLPAP	233			
Qy	119	STERQTRDVRMTOLEVD-----KPTFKDVRSRFGKFKYTKINKTLPPLPPRAITL	170			
Db	234	SIDRSTKPLDRLSLALDRPFFILGKKPFPSPAPLGRHLPKIKQPLP---PAMDR	290			
Qy	171	PKYQPLPPAPPSSSAYFAPKPTFFEVQRPQRSADFSRVLGAEESSHQTKPSSC	230			
Db	291	HERNERLGPVTR-----KPSVPRHGRGPORRE-----NDSDDVHQRPLPQPSL	334			
Qy	231	PSNQNTQKSPPAIAGS-----SYMPGKHSIQARDHTGSMQHC--PAQRCQAASHSP-	281			
Db	335	PSMSNTTTPSRVQPSKNTFFLAHMGAFS---ESNIGFQQSASLPYPYSGQGNRPPL	391			
Qy	282	-----RMLPYNTNEXKPDPTKPKDEKV-WQNEWYIGYSQAQVEDVLKMNKGDTFLVRD	336			
Db	392	RSEGRNLPLVPN--RPQPPSGEETPLDEEYVSYITRPEAAALRKINQDGTFLVRD	449			
Qy	337	CSYKSAEPVVLVYFVYGNKYVKIRFLESNQFALGTGLRGNEMFDSVSDIETHYTFP	396			



T27877  
hypothetical protein ZK470.5 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T27877  
R:Minx, P.  
submitted to the EMBL Data Library, October 1995  
A:Description: The sequence of *C. elegans* cosmid ZK470.  
A:Reference number: Z20433  
A:Accession: T27877  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-443 <MIN>  
A:Cross-references: UNIPARC:UPI000017BD0E, EMBL:U39651, PIDN:AAA80397.1, CESP:ZK470.5  
C:Genetics:  
A:Gene: CESP:ZK470.5  
A:Introns: 39/3; 83/3; 109/3; 214/3; 270/3; 332/3; 352/3; 416/2

Query Match 6.0%; Score 138.5; DB 2; Length 443;  
Best Local Similarity 28.6%; Pred. No. 0.053;  
Matches 46; Conservative 23; Mismatches 49; Indels 43; Gaps 8;

Qy 277 ASHSRMLPYENTNSENKDPDKDEKDVQWNEWITGYSRQAVEDVLKMKENKDGTFLLVRD 336  
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 280 ASHQD-PAPQYSGNGEIPMEQQP-----WYFGRISRERAEDLL-LHGREGEFLVRD 328  
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Qy 337 CSTKSKAEPYVLVYFYGNKYVNVKI-----RELESNQOPAL---GTGLRG 378  
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 329 -SSGNLQKLVL-----EKLDWTQVQTVAKPGDLSISMRGIERNKHFKVQVNDGLLKIG 381  
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Qy 379 NEMFDSVEDIIEHYTPFILLIDGDKAARRKQCYLTQPLP 419  
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 382 NRIFVDNALINHYTTSPIF-----SSPTEKLFITGPLP 415  
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

RESULT 7  
JC7552  
Shb-like adapter protein, Shf - human  
C:Species: *Homo sapiens* (man)  
C:Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 09-Jul-2004  
C:Accession: JC7552  
R:Lindholm, C.K.; Frantz, J.D.; Shoelson, S.E.; Welsh, M.  
Biochem. Biophys. Res. Commun. 278, 537-543, 2000  
A:Title: Shf, a Shb-like adapter protein, is involved in PDGF-alpha-receptor regulation  
A:Reference number: JC7552; MUID: 20548990; PMID:11095946  
A:Accession: JC7552  
A:Molecule type: mRNA  
A:Residues: 1-480 <LIN>  
A:Cross-references: UNIPROT:Q7M4L6; UNIPARC:UPI000017A517  
C:Comment: This protein, a novel adapter protein, has the roles in PDGF-receptor and cell

```

A;Gene: Shf
C;Keywords: apoptosis

Query Match          5.8%; Score 134.5; DB 2; Length 480;
Best Local Similarity 20.7%; Pred. No. 0.11;
Matches 85; Conservative 55; Mismatches 138; Indels 133; Gaps 19;

Qy      26  KNRSWPSLSAK-GRCRAVLEPLP-----DHRNLAVPGG-EKCSNNNDY 69
Db      135  KGSNWRSTTRLIIRLDRLPRELAIEDYADFPDVGTEGSAGASGAPKYPENDGY 194
Qy      70  EDPEFQLLKAWPSMKILPARPIQSEYADTRYFODMMEAPLL-----PPKASVSTRQ 123
Db      195  MEP-----YEAQNM-----AEIRGSKETATQPLFLDYTPYEPEDGATPEGE 237
Qy      124  ----TRDVRMTQ-----LEEVDKP-TPKDVRSQRFKFKYTKINKTLPPLP-----RPA 167
Db      238  GAPWPRSRLPEDDERPPEEVDQPWENKGERISKAFVDIKVIXDLPPWPPPGVGLDSSPS 297
Qy      168  I-----TLPKYQPLPPAPPEESSAYFAPKPTTFPEVQGRQRSAKDFSRVLGAEEESH 222
Db      298  LPDGDRIISGPASLPPEPSLEDSQAQF-----SSSSSSSSSSSSSSSSSSSSSSSSSS 324

```

324



Query Match	5.7%	Score 132;	DB 2;	Length 1603;
Best Local Similarity	21.2%;	Pred.No. 0.73;		
Matches	80; Conservative	51; Mismatched	148; Indels	98; Gaps 18

  

QY	26	KNRSWPSSLSSAKGRCR-----AVLEPLPDHRRNLAGVPG----	GEKCNNSNDYEDPEFQL	76
DB	641	KSSSSSSSGGKRKSPSLTPVTIRTIMSPSGVSTLSPRTVSIGA	FSDPKSEF	700
		.: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .	.: .: .: .: .: .: .: .	
QY	77	LKAWP-----SMKILPARP-IQSEYADTR--YFQD--MMEAPLLLLPPK	115	
DB	701	LKGFALKPICKVBESPRTLNNRAITPPSPVOQSASPKSGKNLLDSILMKPSCMPK	760	
		.: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .	.: .: .: .: .: .: .: .	
QY	116	ASVSTERQTRD---VRMTOLEVDKPT--FK--DVRSQRFGFKYTKINKTLP	PPR-R	165
DB	761	SIASSKKKEPVKANVKQKLSPPLTVDFKIRLPVTNGNSSGTASP	KIEKLPFPAPK	820
		.: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .	.: .: .: .: .: .: .: .	
QY	166	PAITLPKKYOPL-----PPAPPESAA--YFAP-----	RPTTFEVR-GPRQRS	207
DB	821	PMPLAPKLOFSAOFAFPSPPIHHAGVONSAGRNTPIAKRYOFLIPKASRPNFANIP	880	
		.: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .	.: .: .: .: .: .: .: .	

Query Match	5.6%;	Score 130.5;	DB 2;	Length 724;
Best Local Similarity	20.24;	Pred. No. 0.34;		
Matches	72;	Conservative	43;	Mismatches 114;
			Indels 127;	Gaps 15
Qy	148	FKGFKYTKINKTLP	PPRPALTLPKKYQPLPPAP	-----PESSA-----YEAPK 192
Db	69	FPGTVEYVYGRKKISPTPK	---PRPRPLVPVAPGPKSTEAD	SBQOASTLPDLASQFAPP 125
Qy	193	PTPEPQVGRQRSAKDFSR	VLGAEBESHHTQKPESS	-----CPSSNQTQ----- 238
Db	126	DVAP-----PLLKLVEA	IEKKGLECSTLYRTQSSNPAEL	RQLLDCOTASLDLEMFVH 180
Qy	239	-----KSP-PAIASSYV	WPKXHSIQARD-----	HTGSMQHCPAORC 273
Db	181	VLADAFKRYLLDLFPN	PIPVAVSSSELISLAPEVQSBEY	IQLKKLIRSPSIPHQYWLTL 240
Qy	274	QAAASH-----	SRMLPYENTNSK-----	----- 293
Db	241	QYLLKHFFKLQSTSSKNLL	NARVLSELFPLFRPPAASSNT	TEHLIKIEILISTEWNE 300
Qy	294	-----PDPTKP-----	DEKDVWQWNEYIGEYSRQAVED	VLMMKENKDGTFLLVR 336
Db	301	RQPAPALPPKPPKPTT	VANNGMNMSLQDAEWYWGDIS	REEVNEKL-RDTADGTFLLVR 359
Qy	337	CSTKSKRAEPVYLVFY	GNKYVNVKIRFLESNQQFALGT	GLRGNEMFDSVEDIEHY 392
Db	360	ASTKMGD-YTLTLR	KGGN-NKLIKIFHRDKGYG	FDPL-----TFNSVVELINHY 408

RESULT 13  
I49444  
SH3 binding protein - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004  
C:Accession: I49444  
R:Ren, R.; Mayer, B.J.; Cicchetti, P.; Baltimore, D.  
Science 259, 1157-1161, 1993  
A:Title: Identification of a ten-amino acid proline-rich SH3 binding site.  
A:Reference number: I49444; MUID:93174278; PMID:8438166

A;Accession: I49444  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-559 <RES>  
A;Cross-references: UNIPROT:Q06649; UNIPARC:UPI0000027A0E; GB:L14543; NID:g293267; PIDN:  
F;25-128/Domain: pleckstrin repeat homology <PLK>  
F;201-210/Region: proline-rich SH3 binding

Query Match 5.6%; Score 130; DB 2; Length 559;  
Best Local Similarity 21.4%; Pred. No. 0.27;  
Matches 87; Conservative 42; Mismatches 145; Indels 132; Gaps 19;

QY 86 LPARPQESY----ADTRYFQDMMEAPLLI-----PPKASVSTERQTRDVRMT 130  
DB LSSYPMDNEDYEHEDSDSLEPDSGPGMKLEDAITYPPAYPPPPVPPR-----213  
QY 131 QLEBVDKPTKDVRSRFGKFKYTKINKTP-LPPRPAPITLP-----KKY 174  
DB LSSYPMDNEDYEHEDSDSLEPDSGPGMKLEDAITYPPAYPPPPVPPR-----213  
QY 214 -----KPAFSDLP---RAHSFTSKSPSPLLPPPPPKRGLPDTGSAPEDAKDALGLRRV 263  
QY 175 QP---LPPAPPESAYFAFKPTFPEVQRPRORS-----A 207  
DB EPLGLRPATPRRMSDPFMSNVPTPNLRKHPCFRDSVNPGLPFWPTFGHGTSSVSSSTTWA 323  
QY 208 KDFSRLVGAEESHHTK--PESSCPSNONTOK-----SPPAIASSVMPGKHSIQ 257  
DB VATSRNCDKLKSFHLSRGSGPTSEPPVPANKFKFLKIAEPPSPREAAKFAVPV---PVA 380  
QY 258 ARDHTGSMQHCAPQRCQAASHSPRLMPYENTNSEKPD-----PTKPD--- 300  
DB PRPPVQKMP-MPEATVRPVLPRPENTPLPHLQRSPPDGSGRFGSFEPKARQPSQADTGE 439  
QY 301 -----EKDVQNEWYIGYSRQAVEDVLMK-----ENKDGTFVLVDCSTKSKAEPYVL 348  
DB EDSDEYKVPVLPNSVFVNTTESCEVERLPKATDPRGEQDGLYCIKRNSTKSGK---VL 496  
QY 349 VVFYGNKYVNVKIRFLESNQOQFALGTGLRGNEFDSVEDIIEHY 392  
DB VVWDESNKVRNRYI--FEKDSKPY----LEGEVLFASVGSVMVEHY 536

RESULT 14  
A38747  
Phosphatidylinositol 3-kinase (EC 2.7.1.-) 85K chain - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 03-Aug-1992 #sequence\_revision 03-Aug-1992 #text\_change 09-Jul-2004  
C;Accession: A38747  
R;Racobedo, J.A.; Navankasattusas, S.; Kavanagh, W.M.; Milfay, D.; Fried, V.A.; William  
Cell 65, 75-82, 1991  
A;Title: cDNA cloning of a novel 85 kd protein that has SH2 domains and regulates bindin  
A;Reference number: A38747; MUID:91191564; PMID:1849460  
A;Accession: A38747  
A;Molecule type: mRNA  
A;Residues: 1-724 <ESC>  
A;Cross-references: UNIPROT:P70304; UNIPARC:UPI000017C725; GB:M60651  
C;Comment: This protein binds a phosphotyrosine-containing sequence of ligand-activated  
phosphatidylinositol at position 3 of the inositol ring.  
C;Keywords: phosphotransferase  
F;333-428/Domain: SH2 homology <SH2A>  
F;624-718/Domain: SH2 homology <SH2>

Query Match 5.6%; Score 130; DB 2; Length 724;  
Best Local Similarity 22.0%; Pred. No. 0.37;  
Matches 89; Conservative 46; Mismatches 147; Indels 122; Gaps 19;

QY 38 GRCRAVLEPLDHR--RNLAGVPGGKCNNDYED-PEFQLLKWPMSMKILPARIQES 94  
DB GRKX-ISPPTPKRPRPLPVARGSSKTEADTQQALPLPDIAEQFAPDPVAPLLIKLL 136  
QY 95 FYADTRYFQDMMEAPILLPPKASVSTERQTR-----DVRMTQLEEVDKPTFKDVRSORFK 149  
DB EATEKK-----GLECSTLYRTQSS-SNPAELRQLDCCAAVDLEMDIVHVLADA---FK 187

QY 150 GFKYTKINKTFLPP-----PRPAITLPKKYQQLPPAPPE-ESSAYFAPK 192  
DB --RYLADLPNVPVAVYNNMMSLAQELQSPEDCIQLLKLIRLNPINPHOCWUTLQLLLK 245  
QY 193 PTFPEVQRGRQRSKAD-----FSRVL-----GAEESHHTQTKPESSCPSSNQNT 237  
DB 246 HFFK-----LSQASSKNLLNARVLSEIFSPVLPFRPPAASDNTLHLKATLILISTEWNE 300  
QY 238 QKSPPPAIASSSYNPGKHSIQARDHTGSMQHCAPQRCQAASHSPRLMPYENTNSEKPDPT 297  
DB 301 RQAPALP-----PPKPP 312  
QY 298 KP-----DEKDVQNEWYIGYSRQAVEDVLMKENKDGTFVLVDCSTKSKAEPYVL 348  
DB 313 KPTTVANNMNNMNSQDAEWYMGDISREEVNKKL-RDTADGTFVLVDCSTKSKAEPYVL 370  
QY 349 VVFYGNKYVNVKIRFLESNQOQFALGTGLRGNEFDSVEDIIEHY 392  
DB 371 TLRKGGN--NKLKIFHRDQKYGFSDDL---TFNSVVELINHY 408

RESULT 15  
S68957  
adhesive plaque protein precursor - Mediterranean mussel  
C;Species: Mytilus galloprovincialis (Mediterranean mussel)  
C;Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004  
C;Accession: S68957  
R;Inoue, K.; Takeuchi, Y.; Miki, D.; Odo, S.; Harayama, S.; Waite, J.H.  
Submitted to the EMBL Data Library, August 1995  
A;Description: Cloning, sequencing and sites of expression of genes for the hydroxyarginin  
A;Reference number: S68957  
A;Accession: S68957  
A;Molecule type: mRNA  
A;Residues: 1-751 <INO>  
A;Cross-references: UNIPROT:Q27409; UNIPARC:UPI000012AB79; EMBL:D63778; NID:g961463; PIDN:  
C;Keywords: hydroxyproline  
F;1-24/Domain: signal sequence #status predicted <SIG>  
F;45-751/Product: adhesive plaque protein #status predicted <MAT>  
F;123,127,133,173,177,183,227,267,283,293,297,303,307,313,323,327,333,337,343,347,367,373;  
33/Modified site: 4-hydroxyproline (Pro) #status predicted  
F;125,129,135,139,175,179,185,189,225,229,265,269,285,289,295,299,305,309,315,319,325,329;  
79/Modified site: 3',4'-dihydroxyphenylalanine (Tyr) #status predicted  
F;126,136,176,186,226,266,286,296,306,316,326,336,346,366,376,386,406,416,426,436,446,456;  
ans-2,3-cis-3,4-dihydroxyproline (Pro) #status predicted  
F;485,489,495,499,505,509,515,519,525,529,535,539,545,549,555,559,565,569,575,579,585,589;  
dihydroxyphenylalanine (Tyr) #status predicted  
F;537,543,547,553,567,573,577,583,587,593,597,603,607,613,617,623,627,633,637,643,647,653;

Query Match 5.6%; Score 129; DB 2; Length 751;  
Best Local Similarity 23.4%; Pred. No. 0.45;  
Matches 55; Conservative 34; Mismatches 108; Indels 38; Gaps 9;

QY 84 KILPAPRIQESYADTRY---FQDMMEAPILLPPKASVSTERQTRDVRMTQLEBVDKPTF 140  
DB KLSSYKPIKTYNAKNTYPPVYKPKTYTPPYKPKSPYPTYK-----PKPSY 135  
QY 141 KDVRSORFGFKYTKINKTPLPPLPPRAITLPKKYQQLPPAPPESSAYFAPKPTFFBQVR 200  
DB 136 PATYKSKSYPSYKPKTKTYPPTYKPKLTYPPTYKPKPSYPPT-----YKPKPSYPATYK 190  
QY 201 G-----PRORSKDFSRVLGAEESSHHTQKPESSCPSSNONTQKSP-----AIASSSYM 250  
DB 191 SKSSYPSPYKTKTYTPSSYKPKTKTYTPYKPKSYTPYTKSKSYPIYTKASYPSSYK 250  
QY 251 PGKHSIQARDHTGSMQHCAPQRCQAASHSPRLMP-YENTNSEKPD--PT---KP 299  
DB 251 PKK--TYPSTYKPKISYPPYTKAKSPYPTSYRAKPSYSPYTKAKPSYPTPYTKAKP 303

Search completed: March 22, 2006, 01:36:41  
Job time : 43 secs



Result No.	Query	Score	Match	Length	DB	ID	Description
1	23316	99.0	0.0	435	2	Q9JMJ3	MUSE
	2	2305	99.5	435	2	Q9OZE2	MUSE
	3	2070	89.4	392	2	Q8C479	MUSE
	4	1306.5	56.4	428	2	Q7Z7G1	HUMAN
	5	1132	48.9	376	2	Q9P2U9	HUMAN
	6	412.5	17.8	533	1	LCP2	MUSE
	7	410.5	17.7	533	2	Q922M0	MUSE
	8	410	17.7	525	2	Q920L0	RAT
	9	409	17.7	534	2	Q5SUP7	MUSE
	10	398	17.2	533	1	LCP2	HUMAN
	11	398	17.2	533	2	Q53XV4	HUMAN
	12	381	16.5	530	2	Q9DG07	CHICK
	13	360	15.5	501	2	Q5U3P8	BRARE
	14	359	15.5	508	2	Q6GP30	XENLA
	15	355	15.3	501	2	Q8JI36	BRARE
	16	331	14.3	523	2	Q8CASI	MUSE
	17	329.5	14.2	524	2	Q5SUP8	MUSE
	18	319	13.8	456	2	Q75498	HUMAN
	19	319	13.8	456	2	Q8WV28	HUMAN
	20	315.5	13.6	552	2	Q9YGC1	CHICK
	21	309.5	13.4	433	2	Q75499	HUMAN
	22	308	13.3	544	2	Q5HZD8	XENTR
	23	307.5	13.3	457	2	Q4KM52	RAT
	24	304.5	13.1	529	2	Q8JG66	XENLA
	25	299	12.9	549	2	Q5U4V8	XENLA
	26	297.5	12.8	457	2	Q9U0N3	MUSE
	27	295.5	12.8	490	2	Q8JI35	BRARE
	28	287.5	12.4	457	2	Q8E504	MUSE
	29	219.5	9.5	297	2	Q9PD43	MUSE
	30	198.5	8.4	175	2	Q7Z4S9	HUMAN
	31	184.5	8.1	516	2	Q54737	MUSE

```

QY 241 PPAIASSYMPGKHSIQARDHTGSMQHCPCPAQRCAAAASHSPRLPYENTNNTSEKPDPTKPD 300
Db 241 PPAIASSYMPGKHSIQARDHTGSMQHCPCPAQRCAAAASHSPRLPYENTNNTSEKPDPTKPD 300
QY 301 EKDVQNEWYIGYSQAQVEDVLKMKNDGTFLVRDCSTKSKAEPVVLVVFYGNKVNVK 360
Db 301 EKDVQNEWYIGYSQAQVEDVLKMKNDGTFLVRDCSTKSKAEPVVLVVFYGNKVNVK 360
QY 361 IRFLESNQFALGTGLRGNEMFSDVEDIIEHYTYFFILLIDGDKDAARRKQCYLTQPLPL 420
Db 361 IRFLESNQFALGTGLRGNEMFSDVEDIIEHYTYFFILLIDGDKDAARRKQCYLTQPLPL 420
QY 421 ARLLLTQYSSQALHE 435
Db 421 ARLLLTQYSSQALHE 435

RESULT 2
Q9QZE2_MOUSE
ID Q9QZE2_MOUSE PRELIMINARY; PRT; 435 AA.
AC Q9QZE2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CLNK.
GN Names=Clnk;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6;
RX MEDLINE=20029811; PubMed=10562326; DOI=10.1084/jem.190.10.1527;
RA Cao M.Y., Davidson D., Yu J., Latour S., Veilleux A.;
RT "Clnk, a novel SLP-76-related adaptor molecule expressed in cytokine-
stimulated hemopoietic cells.";
RL J. Exp. Med. 190:1527-1534(1999).
DR EMBL; AF187819; AAF14299.1; -; mRNA.
DR HSP; P00524; INZL.
DR Ensembl; ENSMUSG0000039315; Mus musculus.
DR MGI; MGI:1351466; Clnk.
DR GO; GO:0005515; F:protein binding; IPI.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. .; IDA.
DR InterPro; IPR000980; SH2.
DR Pfam; PF00017; SH2; 1.
DR PRINTS; PR00401; SH2DOMAIN.
DR ProDom; PD000093; SH2; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PS50001; SH2; 1.
SQ SEQUENCE 435 AA; 49492 MW; 5CD27EC971FC0EA5 CRC64;

Query Match 99.5%; Score 2305; DB 2; Length 435;
Best Local Similarity 99.5%; Pred. No. 1.6e-133;
Matches 433; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 1 MTSQGNKRTTKEGFDLRFQNVSLKKNRWPSSLSSAKGRCRAVLEPLPDHRRNLAVPGG 60
Db- 1 MTSQGNKRTTKEGFDLRFQNVSLKKNRWPSSLSSAKGRCRAVLEPLPDHRRNLAVPGG 60
QY 61 EKCSNNYDDEPFDQLLKAWPSMKILPARPIQSEYADTRYFQDMWEAPLLLPKASVST 120
Db 61 EKCSNNYDDEPFDQLLKAWPSMKILPARPIQSEYADTRYFQDMWEAPLLLPKASVST 120
QY 121 ERQTRVMTQLEVDKPTFKDVSQRPFKTKTKNTKPLPPRAITLPKKYQPLPPA 180
Db 121 ERQTRVMTQLEVDKPTFKDVSQRPFKTKTKNTKPLPPRAITLPKKYQPLPPA 180
QY 181 PPESAYFAPKPTTFPEVQGRPRQSAKOPSRVLGAEEESHQTKPESSCPSSNQNTQS 240
Db 181 PPESAYFAPKPTTFPEVQGRPRQSAKOPSRVLGAEEESHQTKPESSCPSSNQNTQS 240

```

```

RESULT 3
Q8C479_MOUSE
ID Q8C479_MOUSE PRELIMINARY; PRT; 392 AA.
AC Q8C479;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Mus musculus ES cells cDNA, RIKEN full-length enriched library, linker,
clone:C330034F12 product:cytokine-dependent hematopoietic cell, linker,
full insert sequence. (Fragment).
GN Names=Clnk;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J;
RX MEDLINE=99279283; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
Fleischmann W., Gaasterland T., Giesi C., King B., Kochiwa H.,
Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mommaerts P.,
Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohzuki S.,
Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J;
RA The FANTOM Consortium,
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).

```

```

[4]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RL prepare full-length cDNA libraries for rapid discovery of new genes.";
RN Genome Res. 10:1617-1630(2000).
[5]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsunoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RL sequencing pipeline with 384 multicapillary sequencer.";
RN Genome Res. 10:1757-1771(2000).
[6]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Kajoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Nakazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito K., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK082826; BAC38640.1; -, mRNA.
DR HSP; P00524; INZL.
DR MGI; MGI-1351468; Clnk.
DR GO; GO:0005515; F:protein binding; IPI.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . ; IDA.
DR InterPro; IPR000980; SH2.
DR Pfam; PF00017; SH2; 1.
DR PRINTS; PR00401; SH2DOMAIN.
DR ProDom; PD000093; SH2; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PS00001; SH2; 1.
FT NON TER
SQ SEQUENCE 392 AA; 44749 MW; 29B5B6BBBCA75FDD CRC64;

Query Match 89.4%; Score 2070; DB 2; Length 392;
Best Local Similarity 99.2%; Pred. No. 3.8e-119;
Matches 389; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 44 LEPLDHRRLAGVPGGKCNNSNDYEDPEFQLLKAWPSMKILPARPQSEVADTRYFQ 103
Db 1 LEPLDHRRLAGVPGGKCNNSNDYEDPEFQLLKAWPSMKILPARPQSEVADTRYFQ 60

Qy 104 DMEAPLLLPKASVSTQTRDVRMTQLEVDKPTFKDVRQSRFGKFKYTKTKNTPLPP 163
Db 61 DTMEAPLLLPKASVSTQTRDVRMTQLEVDKPTFKDVRQSRFGKFKYTKTKNTPLPP 120

Qy 164 PRPAITLPKKYQPLPPAPPESSAYFAPKPTFFPVQGRQSRSAKDFSRVLGAEEESHQ 223
Db 121 PRPAITLPKKYQPLPPAPPESSAYFAPKPTFFPVQGRQSRSAKDFSRVLGAEEESHQ 180

Qy 224 TKPSSCSPSSNONTQKSPPAIASSYMPGKHISQIARDHTGSMQHCPCQAQRAASHSPRM 283
Db 181 TKPSSCSPSSNONTQKSPPAIASSYMPGKHISQIARDHTGSMQHCPCQAQRAASHSPRM 240

Qy 284 LPYENTNSEXPDPTKPKDEKQVQNEWYIGYSRQAVEDVLMKENKDGFTFLVRDCSTKSA 343

```

```
QY 418 LPLARLL 425
Db 419 LPLTRHL 426

RESULT 5
Q9P2U9 HUMAN PRELIMINARY; PRT; 376 AA.
ID Q9P2U9
AC Q9P2U9
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE MIST (Fragment).
GN Homo sapiens (Human).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20208989; PubMed=10744659; DOI=10.1093/intimm/12.4.573;
RA Goitsuka R., Kanazashi H., Sasanuma H., Fujimura Y., Hidaka Y.,
RA Tateuno A., Ra C., Hayashi K., Kitamura D.,
RT "A BASH/SLP-76-related adaptor protein Misi/Clnk involved in IgE
RT receptor-mediated mast cell degranulation."
RL Int. Immunol. 12:573-580(2000).
DR EMBL; AB032369; BAA96241.1; -; mRNA.
DR HSSP; Q06124; 2SHP.
DR GO; GO:0005622; C:intracellular; NAS.
DR GO; GO:0005070; F:SH3/SH2 adaptor activity; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR GO; GO:0007242; P:intracellular signaling cascade; IDA.
DR InterPro; IPR000980; SH2.
DR PRINTS; PR00401; SH2DOMAIN.
DR ProDom; PD000093; SH2; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PS50001; SH2; 1.
FT NON TER 1
FT NON TER 376
SQ SEQUENCE 376 AA; 43542 MW; 389421B629B028B2 CRC64;

Query Match 48.9%; Score 1132; DB 2; Length 376;
Best Local Similarity 60.9%; Pred. No. 1.3e-61;
Matches 229; Conservative 42; Mismatches 101; Indels 4; Gaps 3;

QY 19 FQNVSLKNSWPSLSAKGRCRAVLEPLDPHRLNLAGVPGGKCNNDYEDPEQLLK 78
Db 1 FQNFSLPKNSWPRINSATGQYQRMKPLDWRNFAALDGAHGSHDDYDPELRMBE 60

QY 79 AWPMSKILPARPTQSEYADTRFYQDMWEAPLLPPKASVSTRQDRVRMTQLEBVDKP 138
Db 61 TWQSIKILPARPIKESEYADTHYFKVAMDTPLDTRTISISIQPTWNTQ-TELEVRDVP 119

QY 139 TFKDVRSQRFKPKYTKINTPLPPRPAILTPKKYQPLPPAPPESSAYFAPKPTFPEV 198
Db 120 ISRDVRSQNIKGDASVRKNKIPLPPLRPPLITLPKKYQPLPP-EPESRRPPLSQRHTFPEV 178

QY 199 QRGPRORSADFSRVLGAREESHQTKPSSSCSSNQTOKSPATASSYMPGKHSIOA 258
Db 179 QGPFQSISURDLSEVLAEKVPINQRPSTHLENNQNTQEIPLAISSSFTTSNHSVN 238

QY 259 RDHTGSMQHCAPQRCQAAASHSP--RMLPYENTNSKPDPTKPEKDVQMNQNEYIGEYSR 316
Db 239 RDHRGGMQPCSPQRCQPPASCSPHENILPKYTSWRPPPKRSDRKDVQMNQNEYIGEYSR 298

QY 317 QAVEDVLMKENKDGTFIVRDCSTKSKAEPYVLVVFYGNKYNVKIRPLESNQQPALTGL 376
Db 299 QAVEEAFFMKNKDGSLVRDCSTKSKPEPYVLAVFYENKYNVYKIRFLERNQQPALTGL 358

QY 377 RGNEMFDSVEDIEHY 392
Db 379 RGNEMFDSVEDIEHY 374

RESULT 6
LCP2 MOUSE
ID LCP2_MOUSE
AC Q60787;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Lymphocyte cytosolic protein 2 (SH2 domain-containing leucocyte
DE protein of 76 kDa) (SLP-76 tyrosine phosphoprotein) (SLP76).
GN Names=Lcp2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX TISSUE=T-cell lymphoma;
RX MEDLINE=95221345; PubMed=7706237; DOI=10.1074/jbc.270.13.7029;
RA Jackman J.K., Motto D.G., Sun Q., Tanemoto M., Turck C.W., Peltz G.A.,
RA Koretzky G.A., Findell P.R.;
RT "Molecular cloning of SLP-76, a 76-kDa tyrosine phosphoprotein
RT associated with Grb2 in T cells."
RL J. Biol. Chem. 270:7029-7032(1995).
RN [2]
RP INTERACTION WITH SLA.
RX MEDLINE=20130290; PubMed=10662792; DOI=10.1084/jem.191.3.463;
RA Sosinowski T., Pandey A., Dixit V.M., Weiss A.;
RT "Src-like adaptor protein (SLAP) is a negative regulator of T cell
RT receptor signaling."
RL J. Exp. Med. 191:463-474(2000).
CC -!- FUNCTION: Involved in T cell antigen receptor mediated signaling.
CC -!- SUBUNIT: Interacts with SLA. Interacts with the adapter proteins
CC GRB2 and Fyb.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- TISSUE SPECIFICITY: Highly expressed in spleen, thymus, and
CC peripheral blood leukocytes.
CC -!- PTM: Phosphorylated after T-cell receptor activation (By
CC similarity).
CC -!- SIMILARITY: Contains 1 SH2 domain.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; U20159; AAC52189.1; -; mRNA.
CC PIR; B56110; B56110.
CC PDB; 1ORB; X-ray; C/D=231-243.
CC Ensembl; ENSMUSG0000002699; Mus musculus.
CC MGI; MGI:1321402; Lcp2.
CC GO; GO:0050663; P:cytokine secretion; IMP.
CC GO; GO:0045576; P:maat cell activation; IMP.
CC InterPro; IPR001660; SAM.
CC InterPro; IPR011510; SAM 2.
CC InterPro; IPR000980; SH2.
CC Pfam; PF07647; SAM_2; 1.
CC Pfam; PF00017; SH2; 1.
CC PRINTS; PR00401; SH2DOMAIN.
CC ProDom; PD000093; SH2; 1.
CC SMART; SM00454; SAM; 1.
CC SMART; SM00252; SH2; 1.
CC PROSITE; PS50001; SH2; 1.
KW 3D-structure; Phosphorylation; SH2 domain.
FT DOMAIN 422 SH2.
SQ SEQUENCE 533 AA; 60228 MW; 6DDB77782C2E60C3 CRC64;

Query Match 17.8%; Score 412.5; DB 1; Length 533;
```

Best Local Similarity 29.7%; Pred. No. 2.9e-17;			
Matches 148; Conservative 51; Mismatches 175; Indels 125; Gaps 20;			
QY	14	FGDLRFQNVSL-----KNRSPWLSAKGRCRAVLPLPDRHRLAG-----	56
DB	57	FPKLRMPLLSKLSQDINKNEERRIFTRKPOIPRFLBETESHEEDDGGWSSFDDEYSPN	116
QY	57	--VPGGKCNNDYEDP--EFQLL-----KAWPS-----MKILPAPIQ--ESEY	96
DB	117	DDPDGGE---DDGJYESFNEEQALVDDADYEPPPSNNEBALQSSILPPNSFHNTNSMY	173
QY	97	ADTRYFQDMMEAPLLP--PKASV-----	118
DB	174	IDRPPTGKVSQPPVPPPLPRKALPPLPTGRNHSPLSPHPNHEEPSRGNKTKAKLPAP	233
QY	119	STERQTRDVRMTQLEVD-----KPTFXDVRASQRFKGYTKINKTLPPLPPRPATIL	170
DB	234	SIDSTRKPLDLRLAPLDREPFILGKPKPPSPDKPSAPLGRHLPKIQKPLP---PAMDR	290
QY	171	PKYQIPLPPAPPESSAYFAPKPTFPFVQVRQPRORSKDFSRVLGAEESHQTKPSSC	230
DB	291	HERNERLGPVTR-----KPSVPRHGRGPDRE-----NDEDDVHQRLPQPSP	334
QY	231	PSSNQNTQKSPATASS-----SYMPGCKHSIQARDHTGSMQHC--PAQRQQAASHSP-	281
DB	335	PSMSNTTFPSRSVQSPSSKNTFPLAHMPGAFS---ESNIGFOQSASLPPYFSQGGNRPPL	391
QY	282	---RMLPYENTNSEKPDPTKPDKDV--WQNEWIGESRQAVEDVLNKENKGTGFLVRD	336
DB	392	RSEGRNPLPVPN--RQPPSPGSEETPLDEWTVSYITREAEALRKINQDGTFLVRD	449
QY	337	CSTKSKAEPYVLVYFVGNKYNVKIRLESNQFALGTGLRGNMFQSDVEDIIEHYTFP	396
DB	450	SSKKTANNPYVLMVLYKDKVYNIQIRYQESQVYLLGTGLRGKEDFLSVSDIIDYFRMP	509
QY	397	ILLIDGDKAARRKQCVLT 415	
DB	510	LLIDGKNRGS-RYQCTLT 527	
RESULT 7			
Q922M0 MOUSE			
ID	Q922M0	MOUSE PRELIMINARY; PRT; 533 AA.	
AC	Q922M0		
DT	01-DEC-2001	(TrEMBLrel. 19, Created)	
DT	01-DEC-2001	(TrEMBLrel. 19, Last sequence update)	
DT	10-MAY-2005	(TrEMBLrel. 30, Last annotation update)	
DE	Lymphocyte cytosolic protein 2 (Mus musculus adult male bone cDNA,		
DE	product:lymphocyte cytosolic library, clone:9830168E14		
DE	Name=Lcp2;		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;		
OC	Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RP	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RC	STRAIN=FVB/N; TISSUE=Mammary tumor. C3;		
RX	MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;		
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,		
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,		
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,		
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,		
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,		
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,		
RA	Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,		
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullighy S.J.,		
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,		
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,		
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,		
RA	Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,		
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,		

RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,	
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,	
RA	Butterfield J.S.N., Krzywinski M.I., Skalek U., Smallos D.E.,	
RA	Scherer A., Schein J.E., Jones S.J.M., Marra M.A.;	
RT	"Generation and initial analysis of more than 15,000 full-length human	
RT	and mouse cDNA sequences.";	
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).	
RN	[2]	
NUCLEOTIDE SEQUENCE.		
RP	STRAIN=FVB/N; TISSUE=Mammary tumor. C3;	
RC	Strausberg R.;	
RL	Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.	
RN	[3]	
NUCLEOTIDE SEQUENCE.		
RP	STRAIN=C57BL/6J; TISSUE=Bone;	
RC	MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;	
RX	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,	
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,	
RA	Aizawa K., Ozawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,	
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,	
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavani T.,	
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,	
RA	Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,	
RA	Schirml L.M., Staehli F., Suzuki R., Tomita M., Wagner L., Washio T.,	
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,	
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,	
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,	
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,	
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,	
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,	
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,	
RA	Suzuki H., Toyooka K., Wang K.H., Weitz T., Whittaker C., Wilming L.,	
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,	
RA	Hayaishizaki Y.;	
RT	"Functional annotation of a full-length mouse cDNA collection.";	
RL	Nature 409:685-690 (2001).	
RN	[5]	
NUCLEOTIDE SEQUENCE.		
RP	STRAIN=C57BL/6J; TISSUE=Bone;	
RC	MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;	
RX	Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,	
RA	Nikaide I., Osato N., Saito R., Suzuki H., Yamana I., Kiyosawa H.,	
RA	Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,	
RA	Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,	
RA	Schirml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,	
RA	Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,	
RA	Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,	
RA	Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,	
RA	Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,	
RA	Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.A.,	
RA	Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,	
RA	Maglott D.R., Maltais K., Marchionni L., McKenzie L., Miki H.,	
RA	Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,	
RA	Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,	
RA	Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,	
RA	Sandelin A., Schneider C., Sempie C.A., Setou M., Shimada K.,	
RA	Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,	
RA	Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,	
RA	Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang L., Yang L.,	
RA	Yuan Z., Zavalon N., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,	
RA	Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,	
RA	Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,	
RA	Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,	
RA	Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,	
RA	Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,	
RA	Birney E., Hayaishizaki Y.;	

RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573 (2002).  
 RN [6]

RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Bone; DOI=10.1101/gr.145100;  
 RX MEDLINE=20499374; PubMed=11042159; PubMed=11076861; DOI=10.1101/gr.145100;  
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
 RA Konno H., Akizawa J., Nishi K., Kiteunai T., Tashiro H., Itoh M.,  
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
 RA Yamanoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
 RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format  
 RT sequencing pipeline with 384 multicapillary sequencer.";  
 RL Genome Res. 10:1617-1630 (2000).  
 RN [7]

RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Bone;  
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
 RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,  
 RA Konno H., Akizawa J., Nishi K., Kiteunai T., Tashiro H., Itoh M.,  
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
 RA Yamanoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
 RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format  
 RT sequencing pipeline with 384 multicapillary sequencer.";  
 RL Genome Res. 10:1757-1771 (2000).  
 RN [8]

RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Bone;  
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,  
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,  
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,  
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,  
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,  
 RA Saito R., Saichou H., Sakai C., Sakai K., Sakazume N., Sano H.,  
 RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,  
 RA Tagawa A., Takahashi F., Takaku-Akashira S., Takeda Y., Tanaka T.,  
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC006948; AAH06948.1; -; mRNA.  
 DR EMBL: AK036727; BAC29553.1; -; mRNA.  
 DR HSSP: Q60787; IOEB.  
 DR Ensembl: ENSMUSG0000002699; Mus musculus.  
 DR MGI: MGI:1321402; Lcp2.  
 DR GO: GO:0050663; P:cytokine secretion; IMP.  
 DR GO: GO:0045576; P:mast cell activation; IMP.  
 DR InterPro: IPR001660; SAM.  
 DR InterPro: IPR011510; SAM\_2.  
 DR InterPro: IPR000980; SH2.  
 DR Pfam: PF07647; SAM\_2; 1.  
 DR Pfam: PF00017; SH2; 1.  
 DR PRINTS: PR00401; SH2DOMAIN.  
 DR ProDom: PD000093; SH2; 1.  
 DR SMART: SM00454; SAM; 1.  
 DR SMART: SM00252; SH2; 1.  
 DR PROSITE: PSS0001; SH2; 1.  
 DR SEQUENCE 533 AA; 60238 MW; 50A8A025EF0DAD01 CRC64;

Query Match 17.7%; Score 410.5; DB 2; Length 533;  
 Best Local Similarity 29.7%; Pred. No. 3.9e-17;  
 Matches 148; Conservative 50; Mismatches 176; Indels 125; Gaps 20;  
 QY 14 FGDLRFQNVSL-----KNRWSPLSSAKRCRAVLEPLDHRNLAG----- 56  
 DB 57 FPKLRPLLSKLSQDINKNEERSIFTRKQIPRFLEETESHEDDGGWSFFEDYESPN 116  
 QY 57 --VPGKEKSNNDYEDP--EFQLL-----KAWPS-----MKILPARPQ---ESEY 96  
 DB 117 DDDPDGGE---DDGDYSPNEEQALVDDAADYEPSPSNNNEALQSSILPNSPHNTNSMY 173  
 QY 97 ADTRYFQDMMEAPLLLP--PKASV-----EESAYFAPKPTFPEVQRPQRSKADF 118

DB 174 IDRPPTGKVSQPPVPLRPKLPALPPLPTGRNHSPLSPHPNHEBRSRNNKTAKLPAP 233  
 QY 119 STEROTRDRVMTQLEED-----KPTFKVRSQRFKGFYKTKINKTLPPLPPPAITL 170  
 DB 234 SIDRSTKPLDRSLAPLDREPFILGKKPFSDKPSAPLGRHLPKIQKPLP---PAMDR 290  
 QY 171 PKYQPLPPAPPESSAYFAPKPTFPEVQRPQRSKADPSRVLGABEESHQTKPSSC 230  
 DB 291 HERNERLGPVTR-----KPPVPRHGRGPDRE-----NDEDDVHQRPQPQSL 334  
 QY 231 PSSNQNTQKSPPAIASS-----SYMPGKHISIQARDHTGSMQHC--PAQRCQAASHSP- 281  
 DB 335 PSWSSNTFFPSRSVQPSKNTFFPLAHMPGAFS-----ESNIGFQQASLPPYPSQSPGNRPL 391  
 QY 282 ----RMLPYENTNSEKPDPTKPDKDV-WQNEWYIGEYSQAVEDVLMKENKDGTFVLVRD 336  
 DB 392 RSEGRNLPVPV--RPQPSGEEETPLDEWVSVYITRPEAEALRKINQDGTFLVRD 449  
 QY 337 CSTKSKAEPVLVVFVGNKYVVKIRPLESNQOQFALGTGLRGNEMFDSVEDIIEHYTFP 396  
 DB 450 SSKKTANNPVLMVLYKDYNIQIRYQBEESQVYLLGTGLRKEDEFLSVSDIIDYFRMP 509  
 QY 397 ILLIDGKDKAARKQCVLT 415  
 DB 510 LLLIDGNRGS-RYQCTILT 527

# RESULT 8

Q920L0\_RAT PRELIMINARY; PRT; 525 AA.  
 ID Q920L0\_RAT PRELIMINARY;  
 AC Q920L0;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE SLP-76 adaptor protein.  
 GN Name=Lcp2; Synonyms=SLP76;  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Fujii Y., Goithuka G., Wakahara S., Hara T., Nakao T.;  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB072980; BAB71779.1; -; mRNA.  
 DR HSSP: Q60787; IOEB.  
 DR Ensembl: ENSRNOG0000005620; Rattus norvegicus.  
 DR RGD: 619743; Lcp2.  
 DR GO: GO:0007242; P:intracellular signaling cascade; IEA.  
 DR InterPro: IPR011510; SAM\_2.  
 DR InterPro: IPR000980; SH2.  
 DR Pfam: PF07647; SAM\_2; 1.  
 DR Pfam: PF00017; SH2; 1.  
 DR PRINTS: PR00401; SH2DOMAIN.  
 DR ProDom: PD000093; SH2; 1.  
 DR SMART: PSS0001; SH2; 1.  
 DR PROSITE: PSS0001; SH2; 1.  
 DR SEQUENCE 525 AA; 59521 MW; 0B1CDD40DABE7372 CRC64;

Query Match 17.7%; Score 410; DB 2; Length 525;  
 Best Local Similarity 31.9%; Pred. No. 4.1e-17;  
 Matches 127; Conservative 51; Mismatches 148; Indels 72; Gaps 16;  
 QY 68 DYEDPFFQLKAWPSMKILPARPI--QSEYADTRYFQDMMEAP-----LLLP--- 114  
 DB 144 DYEPPSDEEALQN-SILPAKFPNTNSMYIDRPTGKVSQPPVPPQRPMAALPLPT 202  
 QY 115 -----KASVSTERTDRVMTQLEEDVKDTFKDVSQRFGKGYKTKINTPL---- 161  
 DB 203 GRNHSNHEBRSRNNKTAKLPAPSIDRSTKPLD--RSLA-----PPLDEPLLGKK 254  
 QY 162 PPRPRAITLPKKYQP---LPPAPP-----EESAYFAPKPTFPEVQRPQRSKADF 210



```

CC -|- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -|- TISSUE SPECIFICITY: Highly expressed in spleen, thymus, and
CC peripheral blood leukocytes. Highly expressed also in T-cell and
CC monocytic cell lines, expressed at lower level in B-cell lines.
CC Not detected in fibroblast or neuroblastoma cell lines.
CC -|- PTM: Phosphorylated after T-cell receptor activation by ZAP-70.
CC -|- SIMILARITY: Contains 1 SH2 domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; U20158; AAC50135.1; -; mRNA.
CC EMBL; BC016618; AAH16618.1; -; mRNA.
CC PIR; A56110; A56110.
CC PDB; 1H3H; NMR; B=232-241.
CC InAct; Q13094; -.
CC Ensembl; ENSG00000043462; Homo sapiens.
CC HGNC; HGNC:8529; LCP2.
CC H-InvDB; HIX0005403; -.
CC MIM; 601603; -.
CC GO; GO:0005515; F:protein binding; TAS.
CC GO; GO:0006955; P:immune response; TAS.
CC GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . ; TAS.
CC InterPro; IPR001660; SAM.
CC InterPro; IPR011510; SAM_2.
CC InterPro; IPR000980; SH2.
CC Pfam; PF07647; SAM_2; 1.
CC Pfam; PF00017; SH2_1.
CC PRINTS; PR00401; SH2DOMAIN.
CC ProDom; PD000093; SH2; 1.
CC SMART; SM00454; SAM; 1.
CC SMART; SM00252; SH2; 1.
CC PROSITE; PS50001; SH2; 1.
CC 3D-structure; Direct protein sequencing; Phosphorylation; SH2 domain.
CC SH2.
CC DOMAIN 422 530
CC COMPIAS 133 136 Poly-Glu.
CC COMPIAS 198 201 Poly-Pro.
CC SEQUENCE 533 AA; 60188 MW; C5D22F31D36200C8 CRC64;
Query Match 17.2%; Score 398; DB 1; Length 533;
Best Local Similarity 30.2%; Pred. No. 2.3e-16;
Matches 124; Conservative 52; Mismatches 146; Indels 88; Gaps 14;
QY 68 DYEDPEFQLLKAWPSMKILPARPIQSE--YADTRYFQDMMEAP-----LLLPKKA 116
Db 144 DYEPPPSNDEEALQN-SILPAKFPFNSNSMYIDRPPSGKTPQPPVPQRPMAALPPPPA 202
QY 117 -----SVSTEROTRDVRMTQLEVD-----KPT 139
Db 203 GRNHSPLPPQTNEEPPSRNHNKAKLPAPSIDRSTKPLDRSLAPFDREPTLGGKPP 262
QY 140 FKD---VRSRFGKFKYTKINKTLPPLPPPAITLPKYQPLPAPPPESSAYFAPKPTFP 196
Db 263 FSDKPSIPAGRSGLGHLPKIQKPLPP-----TTERHERSSSLP-----GKKPPVP 308
QY 197 EVQGRQPRQSAKDFSVLGAEESESHQTKPE-----SSCPSSNQNTQKSPPAIASSSYM 250
Db 309 KHGWGPDRE-----NDEDDVHORPLQPALLPMSSNTFFSRSTKSPMNPPLSSH 360
QY 251 PGKHSIQARDHTGSMQHCQAQRCQAASHSP-----RMLPYENTNSEKDPPTKPDKDVW 305
Db 361 PGAFS-ESNSFPQSASLPPYFSQGSNRPPIRAEGRNFLPLPNKPRP-PSPAEENSL 418
QY 306 QNEWYIGEYSRQAVEDVLMKENKDGFLVRDCSTKSKAEPPYVLVFGYKNVYNKIRFLE 365
Db 419 NEEWYYSYITRPEAEALRKINQDGFVLVRDSSKTTTNPYVLMLVLYKDKVYNIQIRYQK 478
QY 366 SNQOQFALGTGLRGNEFDSVEDIIHYTFYFPIILLIDGKKAARRKQCYLT 415
Db 479 ESQVYLLGTGLRGKEDFLSVSDIIDYFRKMPPLLIDGKNRGS-RYQCTLT 527

RESULT 12
Q9DG07_CHICK
ID Q9DG07_CHICK PRELIMINARY; PRT; 530 AA.
AC Q9DG07;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE SLP-76 adaptor protein.
GN Name=SLP76;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.

```

```

RESULT 11
Q53XV4_HUMAN
ID Q53XV4_HUMAN PRELIMINARY; PRT; 533 AA.
AC Q53XV4;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE Lymphocyte cytosolic protein 2 (SH2 domain containing leukocyte
DE protein of 76kDa).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Kalnine N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,
RA Koundinya M., Raphael J., Moreira D., Kelley T., Labaer J., Lin Y.,
RA Phelan M., Farmer A.;
RT "Cloning of human full-length CDSs in BD Creator(TM) System Donor
RT vector.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BT007273; AAP35937.1; -; mRNA.
SQ SEQUENCE 533 AA; 60188 MW; C5D22F31D36200C8 CRC64;
Query Match 17.2%; Score 398; DB 2; Length 533;
Best Local Similarity 30.2%; Pred. No. 2.3e-16;
Matches 124; Conservative 52; Mismatches 146; Indels 88; Gaps 14;
QY 68 DYEDPEFQLLKAWPSMKILPARPIQSE--YADTRYFQDMMEAP-----LLLPKKA 116
Db 144 DYEPPPSNDEEALQN-SILPAKFPFNSNSMYIDRPPSGKTPQPPVPQRPMAALPPPPA 202
QY 117 -----SVSTEROTRDVRMTQLEVD-----KPT 139
Db 203 GRNHSPLPPQTNEEPPSRNHNKAKLPAPSIDRSTKPLDRSLAPFDREPTLGGKPP 262
QY 140 FKD---VRSRFGKFKYTKINKTLPPLPPPAITLPKYQPLPAPPPESSAYFAPKPTFP 196
Db 263 FSDKPSIPAGRSGLGHLPKIQKPLPP-----TTERHERSSSLP-----GKKPPVP 308
QY 197 EVQGRQPRQSAKDFSVLGAEESESHQTKPE-----SSCPSSNQNTQKSPPAIASSSYM 250
Db 309 KHGWGPDRE-----NDEDDVHORPLQPALLPMSSNTFFSRSTKSPMNPPLSSH 360
QY 251 PGKHSIQARDHTGSMQHCQAQRCQAASHSP-----RMLPYENTNSEKDPPTKPDKDVW 305
Db 361 PGAFS-ESNSFPQSASLPPYFSQGSNRPPIRAEGRNFLPLPNKPRP-PSPAEENSL 418
QY 306 QNEWYIGEYSRQAVEDVLMKENKDGFLVRDCSTKSKAEPPYVLVFGYKNVYNKIRFLE 365
Db 419 NEEWYYSYITRPEAEALRKINQDGFVLVRDSSKTTTNPYVLMLVLYKDKVYNIQIRYQK 478
QY 366 SNQOQFALGTGLRGNEFDSVEDIIHYTFYFPIILLIDGKKAARRKQCYLT 415
Db 479 ESQVYLLGTGLRGKEDFLSVSDIIDYFRKMPPLLIDGKNRGS-RYQCTLT 527

RESULT 12
Q9DG07_CHICK
ID Q9DG07_CHICK PRELIMINARY; PRT; 530 AA.
AC Q9DG07;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE SLP-76 adaptor protein.
GN Name=SLP76;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.

```



```

OX NCBI_TaxID=9031;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Thymus;
RX MEDLINE=20451093; PubMed=10993915; DOI=10.1084/jem.192.6.847;
RA Ishiai M., Kurosaki M., Inabe K., Chan A.C., Sugamura K., Kurosaki T.;
RT "Involvement of LAT, Gads, and Grb2 in compartmentation of SLP-76 to
RL the plasma membrane.";
DR J. Exp. Med. 192:847-856 (2000).
DR EMBL: AF226988; AAG18493.1; -; mRNA.
DR HSSP: Q60787; 10EB.
DR Ensembl: ENSGALG00000002113; Gallus gallus.
DR GO: GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro: IPR001660; SAM.
DR InterPro: IPR011510; SAM 2.
DR InterPro: IPR000980; SH2.
DR Pfam: PF07647; SAM_2; 1.
DR Pfam: PF00017; SH2_1.
DR PRINTS: PR00401; SH2DOMAIN.
DR PRODOM: PD000093; SH2; 1.
DR SMART: SM00454; SAM; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS00001; SH2; 1.
DR SEQUENCE 530 AA; 59580 MW; CF6CBCEB8D79909A CRC64;

Query Match 16.5%; Score 381; DB 2; Length 530;
Best Local Similarity 29.8%; Pred. No. 2.5e-15;
Matches 123; Conservative 54; Mismatches 151; Indels 85; Gaps 16;

Qy 61 EKCSNND-YEDPFQLKAWPSMKILPARI-QESEVADTRYQDMMEAPLLLPKASV 118
Db 136 EEAHSDSGYEPSPNSNDEAHNV-IPPAKSLANNTYIDRPTSRSSHQPPVPPQPGP 194
Qy 119 STERQTRDVRMTQLEEVDPKTFKDVRSQRFKGYTKINKTPLPPPPPAITL----- 170
Db 195 SPAPASFGGRGASL-----PAFPPPLPGNNDENVK-----PSKPPAPSIDRSTKPLDR 242
Qy 171 -----PKK-----YQPLPPAPPESSAYFAPKPTTPE- 197
Db 243 LGPPFERENPVGRKPGHPEKLLTFLQALGELQAMWPKPVPPSPDRYERGNPSPLRKQI 302
Qy 198 -VQGGPRQSNKDSRVLGAEESH--HQTKPESSCPSSNQNT-----QKSPPAIASSY 249
Db 303 PVKQGWAAQKRPE-----EEEDHIFQRAVQIQLSLPPYSNTSPKSIKAPKPGSNS- 354
Qy 250 MPGRHSIQARDHTGSMQHCPCAPQCOAAASHPRM-----LPYENTNKEKDPPTKPD 301
Db 355 IPGASARSLSASGL---PPRPLGNNSRSPSGTADLRPLIPRQTAHQNTTEDE 411
Qy 302 KDVMQNEWYIGEYRQAVEDVLMKENKDGTFVLVDCSTKSKAEPLYVLVFGYGNKYNVKI 361
Db 412 -DSLNDWEYVAYISRPEAAALRKINKDGTFLVRDSSRKTTHPYLWLVYKDKVYNIQI 470
Qy 362 RFLESNQFALGTGLRGNEFDSVEDIEHYTFPILLIDGKDKAARKQCYL 414
Db 471 RYQEQDQTYLGTGLKGKEDFSSVAHIIDYFQRTPELLIDGKGRGS-RNQCVL 522

RESULT 13
QS03P8 BRARE PRELIMINARY; PRT; 501 AA.
AC QS03P8;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Stp76.
GN ORFNames=zgc:101809;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=955;
RN [1]

```

```

RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Olfactory epithelium;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters K.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RN and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Olfactory epithelium;
RA Director MGC Project;
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC085444; AAH85444.1; -; mRNA.
DR ZFIN: ZDB-GENE-041114-97; ZGC:101809.
DR GO: GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro: IPR000980; SH2.
DR Pfam: PF00017; SH2; 1.
DR PRODOM: PD000093; SH2; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS00001; SH2; 1.
DR SEQUENCE 501 AA; 55995 MW; 5DB1620B9D5D7A90 CRC64;

Query Match 15.5%; Score 360; DB 2; Length 501;
Best Local Similarity 28.3%; Pred. No. 4.5e-14;
Matches 115; Conservative 57; Mismatches 141; Indels 94; Gaps 17;

Qy 61 EKCSNNDYEDPFQLKAWPS---MKILPARIQSEYADTRYQDMMEAPLLLPKAS 117
Db 136 EGSDSNSYEPPESE-----PSEDKAQICPAKPMDCYIDNNRTRVNRSQPPPE- 188
Qy 118 VSTERQTRDVRMTQLEEVDPKTF---KDVRSQRFKGYTKINKTPLPPPPPAITL 174
Db 189 -----RPGGSPSPVPVDRPSVGMNMRREER-----PQTPKRPAPAVDRSKP 232
Qy 175 QPL-----PPA-----PPESSA-----YFAPKPTTPEVQRG 201
Db 233 GTLDRIHPVAVAGRGCTSSLDRAINPPRSVNAAPFRCPAMEAPCDPMRIKPSLPP-SLG 291
Qy 202 PRORSKDF-----SRVLAEEESHQTKPESS--CPSSNQNTQKSPPAIASSYMPGK 253
Db 292 VR-RSASVTPGYSQNRHLDRHEFHDDTARQTSNTFPLHARNPSRPPGTHGQSFQ--- 347
Qy 254 HSIQARDHTGSMQHCPCAPQCOAAASH--SPRMLPYENTNKEKDPPTKDEKDVQNEWYI 311
Db 348 -----TDSVNPSPSLPAKQEAQMTDHRRSARASGPPQDMGGRQD-----MDPAWYL 393
Qy 312 GEYSRQAVEDVLMKENKDGTFVLVDCSTKSKAEPLYVLVFGYGNKYNVKIFLESNQFA 371
Db 394 GQITRGEAESCLRRVNDGTFLVDRSSNRSSNQPYTLVLVYQDKVYNIQIR--RNQDRFM 451
Qy 372 LGTGLRGNEFDSVEDIEHYTFPILLIDGKDK-AARRKQCYLTQP 417
Db 452 LGTGMKASEFERSVDIVNQHKHTPELLIDAKNRESGQKQCALIYP 498

RESULT 14

```



```
Qy 254 HSIQARDHTGSMQHCPAQRCAAAASH--SPRMLPYENTNSEKPDPTKPDDEKDVWQNEWYI 311
Db 348 -----TDSVNFPSRSLPAKLOEAMTDHRRSARASPPQTDMGGRQD-----MDPAWYL 393
Qy 312 GEYSRQAVEDVLMKENKDGTFVVRDCSTKSKAEPYVLVVFYGNKVYNNVKIRFLESNOQFA 371
Db 394 GQITRGEAESCLRRVNRDGFVVRDSSNRFSNQPYTLVVLVYQDKVYNIQIR--RNQDRFM 451
Qy 372 LGTGLRGNEPFDSVEDIIEHYTYPPILLIDGKOK-AAARRKQCYLTQP 417
Db 452 LGTGMKASESFERVSDIVNQHKHTPLLLIDAKNRESGQOKOCALIYP 498
```

Search completed: March 22, 2006, 01:35:55  
Job time : 236 secs

*This Page Blank (Copy)*

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Rup on: March 22, 2006, 01:36:09 ; Search time 46 Seconds  
(without alignments)  
781.824 Million cell updates/sec

Title: US-10-717-619-2

Perfect score: 2316

Sequence: 1 MTSQGNKRTTKEGFGDLRFQ.....QPLPARLLLTQYSSQALHE 435

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/1/iaa/5-COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/6-COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/H-COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/PCTUS-COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/RE-COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2316	100.0	435	2	US-09-856-061-2
2	1316.5	56.8	428	2	US-10-104-047-3344
3	1132	48.9	376	2	US-09-856-061-4
4	398	17.2	533	2	US-09-949-016-6629
5	391	16.9	602	2	US-09-949-016-10221
6	319	13.8	449	2	US-09-355-214-3
7	319	13.8	456	1	US-08-819-013-1
8	319	13.8	456	2	US-09-355-214-1
9	319	13.8	470	2	US-09-949-016-7976
10	287.5	12.4	457	2	US-09-355-214-5
11	213	9.2	84	2	US-09-513-999C-7233
12	165.5	7.1	474	2	US-08-729-416C-1
13	165.5	7.1	474	2	US-09-433-353-1
14	165.5	7.1	594	2	US-08-729-416C-7
15	165.5	7.1	594	2	US-09-433-353-7
16	146	6.3	529	2	US-09-949-016-8918
17	141.5	6.1	474	2	US-08-807-342B-2
18	140	6.0	469	1	US-08-353-550-6
19	140	6.0	469	1	US-08-551-687-6
20	137.5	5.9	474	2	US-08-729-416C-11
21	137.5	5.9	474	2	US-09-433-353-11
22	137.5	5.9	594	2	US-08-729-416C-17
23	137.5	5.9	594	2	US-09-433-353-17
24	137	5.9	143	2	US-10-104-047-3542
25	136.5	5.9	605	2	US-09-949-016-8269
26	136.5	5.9	630	2	US-09-949-016-7270
27	135	5.8	272	2	US-09-270-767-45259

ALIGNMENTS

RESULT 1

US-09-856-061-2

; Sequence 2, Application US/09856061

; Patent No. 6831151

; GENERAL INFORMATION:

; APPLICANT: GOLTISUKA, RYO

; TITLE OF INVENTION: A mast cell-specific adapter molecules and cDNAs thereof

; FILE REFERENCE: 2001-0614A/WMC/00653

; CURRENT APPLICATION NUMBER: US/09/856,061

; PRIOR FILING DATE: 2001-07-10

; PRIOR APPLICATION NUMBER: JP11-263778

; PRIOR FILING DATE: 1999-09-17

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO: 2

; LENGTH: 435

; TYPE: PRT

; ORGANISM: Mus musculus

; US-09-856-061-2

Query Match 100.0%; Score 2316; DB 2; Length 435;  
Best Local Similarity 100.0%; Pred. No. 5.8e-204;  
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MTSQGNKRTTKEGFGDLRFQNVSLLNRSWPSLSAKGRCAVLEPLDHRNLAGVPGG	60
Db	1	MTSQGNKRTTKEGFGDLRFQNVSLLNRSWPSLSAKGRCAVLEPLDHRNLAGVPGG	60
Qy	61	EKCNNDYEDPFFQLLKAWPSMKILPAPRIQSEYADTRYFQDMMEAPLLPPKASVST	120
Db	61	EKCNNDYEDPFFQLLKAWPSMKILPAPRIQSEYADTRYFQDMMEAPLLPPKASVST	120
Qy	121	EROTRDRVMTQLBEVDKPTFKVRSORFGKFKYTKINKTLPDPPRAITLPPKYQPLPPA	180
Db	121	EROTRDRVMTQLBEVDKPTFKVRSORFGKFKYTKINKTLPDPPRAITLPPKYQPLPPA	180
Qy	181	PPESAYFAPKPTFFEVQPGPQRSQAKDFSRVLGAEESHHQTKPSSCPSSNQNTQKS	240
Db	181	PPESAYFAPKPTFFEVQPGPQRSQAKDFSRVLGAEESHHQTKPSSCPSSNQNTQKS	240
Qy	241	PPAIASSYMPGKHSTQARDHTGSMQHCPCQARQAAASHSPRMLPYENTNSEKPDPTKPD	300
Db	241	PPAIASSYMPGKHSTQARDHTGSMQHCPCQARQAAASHSPRMLPYENTNSEKPDPTKPD	300
Qy	301	EKDWQNEWYIGYSRQAVEDVLMKENKQGTFLVRDCSTKSKAEPPYVLVYFGKVNK	360
Db	301	EKDWQNEWYIGYSRQAVEDVLMKENKQGTFLVRDCSTKSKAEPPYVLVYFGKVNK	360
Qy	361	IRFLESNOQALGTGLRGNEFMFSDVETIIHHYTFPILLIDGKDKAARRKOCYLTQPL	420
Db	361	IRFLESNOQALGTGLRGNEFMFSDVETIIHHYTFPILLIDGKDKAARRKOCYLTQPL	420

Db 361 IRFLESNQFALGTGLRGNEFMDSVEDIIHYTFPIILLIDGDKAARRKQCYLTQPLPL 420  
Qy 421 ARLLITQYSSQALHE 435  
Db 421 ARLLITQYSSQALHE 435  
RESULT 2  
US-10-104-047-3344  
; Sequence 3344, Application US/10104047  
; Patent No. 6943241  
; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: No. 6943241el full length cDNA  
; FILE REFERENCE: HI-A0105  
; CURRENT APPLICATION NUMBER: US/10/104,047  
; CURRENT FILING DATE: 2002-03-25  
; PRIOR APPLICATION NUMBER:  
; PRIOR FILING DATE:  
; NUMBER OF SEQ ID NOS: 4096  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3344  
; LENGTH: 428  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-104-047-3344

Query Match 56.8%; Score 1316.5; DB 2; Length 428;  
Best Local Similarity 62.1%; Pred. No. 2.8e-112;  
Matches 266; Conservative 46; Mismatches 111; Indels 5; Gaps 4;  
Qy 1 MTSQGNKRTTKEGFGDLRFQNVSLLNKRSWPSLSAKGRCRAVLPLPDHRRNLGAVPGG 60  
Db 1 MNRQGNKRTTKEGSDNLDKFNFLPKNRSWPRINSATGQYQRMNKPILLDWRNFAAVLDG 60  
Qy 61 EKCSNNNDYDPPFOLLKAWPSMKILPARPIQSEYADTRYFODMMEAPLLPPKASVST 120  
Db 61 AKGHSDDDDYDDPELRMEETWQSIKILPARPIKESEYADTHYFKVAMDTPLDTRTSISI 120  
Qy 121 ERQTRDVRMTQLEEVDKPTFKDVRQRFGKFTYKINKTLPPLPPPAITLPPKYQPLPPA 180  
Db 121 GQPTWNTQ-TRLERVDKPIISKVRSONIKGDASVRKNKIPLPPLPPPLITLPPKYQPLPP- 178  
Qy 181 PPESSAYFAPKPTFFPEVQGRQPRQSAKDFSRVLGABEESHQTKPSSCPSSNQNTQKS 240  
Db 179 EPSSRRPPLUSQRTFFPEVQRMPSQISLRDLSEVLEAEKVPHNQRPSTHLENQNTQEI 238  
Qy 241 PPAIASSYMPGKHSIQARHDTGSMOHCPCAPQRCQAAASHSP--RMLPYENTNSEKPDPTK 298  
Db 239 PLAISSSFTTSHSVQNRDRHGMQPCSPQRCQPAPSCSPHENILPYKTSWRPFPKR 298  
Qy 299 PDEKDVQNWYIGEYSRQAVEDVLKMKENKDGTFVLVDCSTKSKABPPYLVVFGNKNVYN 358  
Db 299 SDRKDVQHNWYIGEYSRQAVEAEAFMKENKDGSLVDCSTKSKKEEPPYVLAVFYENK 358  
Qy 359 VKIRFLESNQFALGTGLRGNEFMDSVEDIIHYTFPIILLIDGKOK-AARRKQCYLTQ 417  
Db 359 VKIRFLESNQFALGTGLRGDEKFDSDIEDIIHYTKNFPPIILLIDGKOKTGVRHQCHLT 418  
Qy 418 LPLARLLL 425  
Db 419 LPLTRHLL 426

RESULT 3  
US-09-856-061-4  
; Sequence 4, Application US/09856061  
; Patent No. 6831151  
; GENERAL INFORMATION:  
; APPLICANT: GOITSUKA, Ryo  
; TITLE OF INVENTION: A mast cell-specific adapter molecules and cDNAs thereof  
; FILE REFERENCE: 2001-0614A/WMC/00553  
; CURRENT APPLICATION NUMBER: US/09/856,061

; CURRENT FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: JP11-263778  
; PRIOR FILING DATE: 1999-09-17  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 376  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-856-061-4  
Query Match 48.9%; Score 1132; DB 2; Length 376;  
Best Local Similarity 60.9%; Pred. No. 2e-95;  
Matches 229; Conservative 42; Mismatches 101; Indels 4; Gaps 3;  
Qy 19 FQNVSLLNKRSWPSLSAKGRCRAVLPLPDHRRNLGAVPGGKCSNNNDYDPEFOLLK 78  
Db 1 FQNFSLPKNRSWPRINSATGQYQRMNKPILLDWRNFAAVLDGAKGHSDDDDYDDPELRME 60  
Qy 79 AWPMSKILPARPIQSEYADTRYFODMMEAPLLPPKASVSTQTRDVRMTQLEEVDPK 138  
Db 61 TWQSIKILPARPIKESEYADTHYFKVAMDTPLDTRTSISIQPTWNTQ-TRLERVDKP 119  
Qy 139 TFKDVRQRFGKFTYKINKTLPPLPPPAITLPPKYQPLPPAPPESSAYFAKPTPEV 198  
Db 120 ISRDVRSQNIKGDASVRKNKIPLPPLPPPLITLPPKYQPLPP-EPSSRRPPLSQRTFFPEV 178  
Qy 199 QRQPRQSAKDFSRVLGABEESHQTKPSSCPSSNQNTQKSPPAIASSYMPGKHSIQ 258  
Db 179 QGMPQSISLRDLSEVLEAEKVPHNQRPSTHLENQNTQEIPLAISSSFTTSHSVQ 238  
Qy 259 RDHTGSMOHCPCAPQRCQAAASHSP--RMLPYENTNSEKPDPTKDEKDVQNWYIGEYS 316  
Db 239 RDHRGGMQPCSPQRCQPAPSCSPHENILPYKTSWRPFPKRSDRDKDVQHNWYIGEYS 298  
Qy 317 QAVEDVLKMKENKDGTFVLVDCSTKSKABPPYLVVFGNKNVYNVKIRFLESNQFALGTGL 376  
Db 299 QAVEEAFMKENKDGSLVDCSTKSKKEEPPYVLAVFYENKYNVKIRFLERNQFALGTGL 358  
Qy 377 RGNEMFDSVEDIIHY 392  
Db 359 RGDEKFDSDIEDIIHY 374

RESULT 4  
US-09-949-016-6629  
; Sequence 6629, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6629  
; LENGTH: 533  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-6629

Query Match 17.2%; Score 398; DB 2; Length 533;  
Best Local Similarity 30.2%; Pred. No. 7.2e-28;  
Matches 124; Conservative 52; Mismatches 146; Indels 88; Gaps 14;

Qy	68	DYDPEFOLLKAWPSMKILPARPIQESE--YADTRYFQDMMEAP-----LLLPKKA	116
		:         :	
Db	144	DYEPPSPNDEBALQN-SILPAKFPFNNSMYIDRPPSGKTPOQPVPVPPORPMAALPPPPA	202
		:         :	
Qy	117	-----SVSTERQTRDVMTQLLEVD-----KPT	139
		:   :   :	
Db	203	GRNHSLPLPPPOTNNHEEPSRSRNHKTAKLPAPSIDRSTKPPLDRSLAPFDREFTLGKKPP	262
Qy	140	FKD---VRSQRFKPGKYTKINKTLPPLPPRPATITLPKKYQPLPAPPESSEAYFAKPCTFP	196
		:         :	
Db	263	FSDXPSIPAGSLGHLPKIQKPLPP-----TTHERSSFLP-----GKKPPVP	308
Qy	197	EVOGRQRORSKADFSRVLGAEESHHTQKPE-----SSCPSSNQNTOKSPAIASSSYM	250
		:         :	
Db	309	KHGWPDDR-----NDBDVHORPLQPALPMSNSTFPSRSTKPSPMNPLPSSHM	360
		:         :	
Qy	251	PKHSIQARDHTGSMQHCPAQRCQAASHSP-----RMLPYENTWSEXPDTPKDPKOVW	305
		:         :	
Db	361	PGAWS-ESNSRFPQSASLPFYFSQGPSNRPRRAEGRNFFPLPNKNRP-PSAAEEENSL	418
Qy	306	QNEWYIGYSQAQVEDVLMMKENCKGTFLVRDCSTKSKEPYVVLVVFYGKVNVNKIRFILE	365
		:         :	
Db	419	NEEWVSYTTPAEAARLKINODGTFLVRDSSKKTTPNYVMVLVKDYNIQIRYQK	478
Qy	366	SNOQFALGTLRGNEMFDSVEDIEHYTYFPIILLIDGXDKAARRKOCYLIT	415
		:   :   :	
Db	479	BSQVYLLGTGLRGEDFLFSVDIIDYFRKMPLLLIDGKNRGS-RYOCTLIT	527
		:         :	
<b>RESULT 5</b>			
US-09-949-016-10221			
; Sequence 10221, Application US/09949016			
; Patent No. 6812339			
; GENERAL INFORMATION:			
; APPLICANT: VENTER, J. Craig et al.			
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED			
; FILE REFERENCE: CL001307			
; CURRENT APPLICATION NUMBER: US/09/949,016			
; PRIOR FILING DATE: 2000-04-14			
; PRIOR FILING DATE: 2000-04-14			
; PRIOR FILING DATE: 2000-10-20			
; PRIOR FILING DATE: 2000-10-20			
; PRIOR FILING DATE: 2000-10-03			
; PRIOR FILING DATE: 2000-10-03			
; PRIOR FILING DATE: 2000-09-08			
; NUMBER OF SEQ ID NOS: 207012			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 10221			
; LENGTH: 602			
; TYPE: PRT			
; ORGANISM: Human			
US-09-949-016-10221			
Query Match 16.9%; Score 391; DB 2; Length 602;			
Best Local Similarity 30.7%; Pred.No. 3.8e-27;			
Matches 124; Conservative 50; Mismatches 154; Indels 76; Gaps 13;			
Qy	68	DYDPEFOLLKAWPSMKILPARPIQESE--YADTRYFQDMMEAP-----LLLPKKA	116
		:         :	
Db	213	DYEPPSPNDEBALQN-SILPAKFPFNNSMYIDRPPSGKTPOQPVPVPPORPMAALPPPPA	271
		:         :	
Qy	117	-----SVSTERQTRDVMTQLLEVD-----KPT	139
		:   :   :	
Db	272	GRNHSLPLPPPOTNNHEEPSRSRNHKTAKLPAPSIDRSTKPPLDRSLAPFDREFTLGKKPP	331
Qy	140	FKD---VRSQRFKPGKYTKINKTLPPLPPRPATITLPKKYQPLPAPPESSEAYFAKPCTFP	196
		:         :	
Db	332	FSDXPSIPAGSLGHLPKIQKPLPP-----TTHERSSFLP-----GKKPPVP	377
Qy	197	EVOGRQRORSKADFSRVLGAEESHHTQKPE-----SSCPSSNQNTOKSPAIASSSYM	256
		:         :	
Db	378	KHGWPDDR-----ENVYSTKLHORPLQPALPMSNSTFPSRSTKPSPMNPLPSSHMGATS-	434
		:         :	

QY 181 PPBESSAYFAPKPTFFVQGRQPRQSAKDFSRVLGAEESHQHTKPSSCPSSNONT--- 237  
DB 176 EDNDENYVHTTESSPPPEKAPM-----VNRSTKPNSSTPASPPGTASG 219  
QY 238 -----QKSPPAIASSYMP--GKH-----SIQARDHTGSM---QHCPAQRCAA- 276  
DB 220 RNSGAWETKSPPP-AAFSPPLPRAGKKTPTPLKTTTPVASQQNASVCEKEKPIPAERHRGSS 278  
QY 277 ----ASHSPRMLPYENTNSKDDP-----TKPDEKDVWQNE 308  
DB 279 HRQEAQVSPVPPAQKIQHQPILPRFTTEGGNPTVDGPLPFSFSSNSTISEQAGVLCKP 338  
QY 309 WYIGEYRQAVEDVLKMKENKDGTFVLVDCSTKSAEYVILVVFYGNKVVNVKIRFLESNQ 368  
DB 339 WYAGACDRKSAEALHRSNKDGSFLIRKSSGHDSKQPYTLVVFNKRKVVNIPIVRFIEATK 398  
QY 369 QFALGTGLRGNEPDSVEDIIHYTYFPILLIDGKD 404  
DB 399 QYALGRKNGEYFGSVAEIRNHQHSPLVLDSQN 434

RESULT 7  
US-08-819-013-1  
; Sequence 1, Application US/08819013  
; Patent No. 5994522  
; GENERAL INFORMATION:  
; APPLICANT: Chan, Andrew C.  
; TITLE OF INVENTION: BLNK PROTEINS  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert  
; STREET: Four Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States  
; ZIP: 94111-4187  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/819,013  
; FILING DATE: 17-MAR-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/788,322  
; FILING DATE: 24-JAN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Silva, Robin M.  
; REGISTRATION NUMBER: 38,304  
; REFERENCE/DOCKET NUMBER: A-64383-1/RFT/RMS  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 781-1989  
; TELEFAX: (415) 398-3249  
; TELEX: 910 277299  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 456 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; US-08-819-013-1

Query Match 13.8%; Score 319; DB 1; Length 456;  
Best Local Similarity 25.0%; Pred. No. 1e-20;  
Matches 99; Conservative 58; Mismatches 139; Indels 100; Gaps 12;

QY 64 NSNNYDEDFQLKAWPSMKILPARIQSEYADTRYFQDMMEAP---LLLPPKASVST 120  
DB 91 NADDSYEPFVPE---QETRPVHPALPFARGEYIDNRSSQ--RHSPFPKTLPSKPSWPS 144

QY 121 ERQTRDVRMTQLEBVDKPTFKDVRQSQRFGKFKYTKINKTLPPLPPPAITLPPKYQPLPPA 180  
DB 145 EKARLSTLTALOKQPV-----PPKPKGLLEDEADYVVPV 182  
QY 181 PPBESSAYFAPKPTFFVQGRQPRQSAKDFSRVLGAEESHQHTKPSSCPSSNONT--- 237  
DB 183 EDNDENYVHTTESSPPPEKAPM-----VNRSTKPNSSTPASPPGTASG 226  
QY 238 -----QKSPPAIASSYMP--GKH-----SIQARDHTGSM---QHCPAQRCAA- 276  
DB 227 RNSGAWETKSPPP-AAFSPPLPRAGKKTPTPLKTTTPVASQQNASVCEKEKPIPAERHRGSS 285  
QY 277 ----ASHSPRMLPYENTNSKDDP-----TKPDEKDVWQNE 308  
DB 286 HRQEAQVSPVPPAQKIQHQPILPRFTTEGGNPTVDGPLPFSFSSNSTISEQAGVLCKP 345  
QY 309 WYIGEYRQAVEDVLKMKENKDGTFVLVDCSTKSAEYVILVVFYGNKVVNVKIRFLESNQ 368  
DB 346 WYAGACDRKSAEALHRSNKDGSFLIRKSSGHDSKQPYTLVVFNKRKVVNIPIVRFIEATK 405  
QY 369 QFALGTGLRGNEPDSVEDIIHYTYFPILLIDGKD 404  
DB 406 QYALGRKNGEYFGSVAEIRNHQHSPLVLDSQN 441

RESULT 8  
US-09-355-214-1  
; Sequence 1, Application US/09355214  
; Patent No. 6638724  
; GENERAL INFORMATION:  
; APPLICANT: Washington University  
; TITLE OF INVENTION: BLNK PROTEINS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert  
; STREET: Four Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States  
; ZIP: 94111-4187  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/355,214  
; FILING DATE: 23-Jul-1999  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/819,013  
; FILING DATE: 17-MAR-1997  
; APPLICATION NUMBER: US 08/788,322  
; FILING DATE: 24-JAN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Silva, Robin M.  
; REGISTRATION NUMBER: 38,304  
; REFERENCE/DOCKET NUMBER: FP-64383-2/RFT/RMS  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 781-1989  
; TELEFAX: (415) 398-3249  
; TELEX: 910 277299  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 456 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-355-214-1

Query Match 13.8%; Score 319; DB 2; Length 456;



```

Qy 238 -----QKSPRAIASSYMP--GKH-----SIOARDHTGSM---OHCPAORCOAA- 276
Db 241 RNSGAWETKSPPP-AAPSPUPRAGKQPTTLPKTPPVASQQNASVCEEKPIPAERHGRSS 299
Qy 277 ----ASHSPMLPYENTNSSEKDDP-----TKPDEKDVWQNE 308
Db 300 HRQEAQVSPVPFPAQKQIHQKPIPLPRFTBGGNPTVDGPLSPSSNSTISEQAGVLCKP 359
Qy 309 WYTGYSRQAVEDVLKMKENKDGTFILVRDCKTSKABPYLVWYGNKYNVVKIRFLESNQ 368
Db 360 WYAGACDRKSAAEALHRSNKGSLIRKSSGCHDSKQPYTLVWFENKRVYNIPVRFIATK 419
Qy 369 OFALGTCLGRNMFDSVEDLIEHYTFYFPIILLIDGK 404
Db 420 OYALGRKKNKEEYFGSVAEILRNHQHSPLVLIDSQN 455

RESULT 10
US-09-355-214-5
; Sequence 5, Application US/09355214
; Patent No. 6638724
; GENERAL INFORMATION:
; APPLICANT: Washington University
; TITLE OF INVENTION: BLANK PROTEINS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESS: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/355,214
; FILING DATE: 23-Jul-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/919,013
; FILING DATE: 17-MAR-1997
; APPLICATION NUMBER: US 08/788,322
; FILING DATE: 24-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: FP-64383-2/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELE: 910 277299
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 457 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-355-214-5

Query Match 12.4%; Score 287.5; DB 2; Length 457;
Best Local Similarity 25.5%; Pred. No. 7.9e-18;
Matches 106; Conservative 53; Mismatches 139; Indels 117; Gaps 16;

Qy 66 NNDYDEPQFOLLKXAPSMKILPAPIOESVADTRYFQDMMEAPLLLPKXASVSTERQTR 125
Db 69 DSDYENPDEH---SDSEMYVNPAAETGDDSYE-----PPPA-----EQQTR 106
Ov 126 DVR-----MTOLIEVDKTFKDVASQRFK-GFKYTKINKT--PLPPPPPAITLP-----KK 173

```

QV 126 DVR-----MTQLEEVDPKPTFKDVRSSQRFK-GFKYTKINKT--PLPPPRAITLP-----KK 173

Db 107 VHPALPFTTGEYVDNR-----SQHSPPFKTLPSKPSWPSAKARLASTLPAPNSLQK 161  
Qy 174 YQPLPAPP---BESSAYFAPKPTFPVQVGRQRS-----AKDFSVLGAEEESHOT 224  
Db 162 PQ-VPPKPKDLLEADYVVEVDNDENYTHPRESSPPPAEKAPMVRSTKPNSSSKHMS 220  
Qy 225 KPES-----SCSSNQNTOKSP-----PAIASSSYMFGKHSIQAR 259  
Db 221 PPGTVAGRNQVWDSKSLPAAPSPUPRAGKATATPLKTPVPPLPNASNVCEKVPFAE 280  
Qy 260 DHTGSMQHCQARQCAASH-----SPRMLPYENTINSEKPDPT----- 297  
Db 281 RHRS-----SHRQDTVQSPVFPPTQKPVHQKPVLPFRPEAGSPAADGPFH 327  
Qy 298 -----KPDEKDVQWQNEWYIGEYSRQAVEDVLMKENKDGTFVLVRDCSTKSKAEPVVLV 349  
Db 328 SFPFNLTFADQEGELLGKPYAGACDKFAEALHRSNKDGSFLIRKSGFHDGSKQPVTLV 387  
Qy 350 VFYGNKYVNVKIRFLESNQFALGTGLRGNEMFDSVEDIIHYTYPIILLIDGKD 404  
Db 388 AFFNKRVINYPVPIEATKQYALGKKNGBEYFGSVVEIVNSHOHNPVLIDSQN 442

## RESULT 11

US-09-513-999C-7233  
; Sequence 7233, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; Patent No. 6783961  
; FILE REFERENCE: 59.US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 7233  
; LENGTH: 84  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: 68  
; OTHER INFORMATION: Xaa=Asp or Glu  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: 69  
; OTHER INFORMATION: Xaa=Ala or Asp or His or Asn or Pro or Ser or Thr or Tyr

US-09-513-999C-7233  
Query Match 9.2%; Score 213; DB 2; Length 84;  
Best Local Similarity 48.8%; Pred. No. 4.6e-12;  
Matches 41; Conservative 17; Mismatches 26; Indels 0; Gaps 0;  
Qy 1 MTSQGNKRTTKEGFDLRFQNVSLKRNPSWPSLSAKGRCAVLEPLPDHRRNLGAVPGG 60  
Db 1 MNRQGNKRTTKESNDLKQNFSLPKNRSWPRINSATGQYQRMNKKLLDWFNFAAVLDG 60  
Qy 61 EKCSNNNDYEDPEFQLLKAWPSMK 84  
Db 61 AKGHSDDXDDPELRMEETWQSIK 84

## RESULT 12

US-08-729-416C-1  
; Sequence 1, Application US/08729416C  
; Patent No. 6013767  
; GENERAL INFORMATION:

; APPLICANT: NAKAMURA, TAKESHI  
; TITLE OF INVENTION: NOVEL BRAIN-SPECIFIC ADAPTER MOLECULE GENE  
; TITLE OF INVENTION: THEREOF, AND ANTIBODY THERETO  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PILLSBURY, MADISON & SUTRO, L.L.P.  
; STREET: 1100 NEW YORK AVENUE, N.W.  
; CITY: WASHINGTON  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005-3918  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/729,416C  
; FILING DATE: 11-OCT-1996  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: PERRY, GLENN J.  
; REGISTRATION NUMBER: 28458  
; REFERENCE/DOCKET NUMBER: 7898/225948  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-861-3000  
; TELEFAX: 202-822-0944  
; TELEX: 6714627 CUSH  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 474 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-729-416C-1

Query Match 7.1%; Score 165.5; DB 2; Length 474;  
Best Local Similarity 23.2%; Pred. No. 1.3e-06;  
Matches 106; Conservative 60; Mismatches 161; Indels 129; Gaps 25;  
Qy 13 GFGLRFQNVSLKRNPSWPSLSAKGRCAVLEPLPDHRRNLGAVPGGCKSNNDY--- 69  
Db 94 GKSNIQFAGMSISLTISTASLNLRTPDSKQII--ANHMKRSISFASGGDP--DTTDVAY 149  
Qy 70 --BDP-----EFQLLKAWPSMKILPARIQESEYADTRYFQDMMEAPLLLPKASVSTERQ 123  
Db 150 VAKDPVNRRAACHILECCDG---LAQDVIGSIGQAFELRFKQYLQCPTKIP---ALHDMQ 203  
Qy 124 TRDVRMTQLEEVDPKPTFKDVRQRFKGYTKINKTLPDPRPAITLPPKKYQPLPPAPPE 183  
Db 204 SLDEPWTE-EEGD-----GSDHPYNSIPSKMPPPGGFLDTLKRPHAP-- 247  
Qy 184 ESSAYFAPK-----PTPPE-VQRGPRQRSKDFSRVLGAEEESHQTKPESSCP 231  
Db 248 -DTAQFAGKEQTYVYQGRHLGDTFGEDWQOTFLRQSSD-----IVSTPEGLHVAPTGEAP 302  
Qy 232 SSNQNTOKSPPAITASSSYMFGKHSIQARDHTGSMQHCQAPQRCQAASHSPR-----MLPY 286  
Db 303 -TYVNTQOIIPP-----QAWPA--AVSSAESSPKDLFDMKPF 336  
Qy 287 ENTNSKPP-----DPTKPDEx---DVMQNEWYIGEYSRQAVEDVLMKE 326  
Db 337 EDALKNQPLGVLVSKAASVECTSPVSPRAPDAKMLELQAEYQEMSKAEAGLL--- 393  
Qy 327 NKDGTFLVRDCSTKSKAEP--YVLVVFYGNKYVNVKIRFLESNQFALGTGLRGNEMFDS 384  
Db 394 EKDGDFLVR---KSTTNPQSVLTGMHNGQAKHLLLDVPE-----GTIRTKDRVDFS 442  
Qy 385 VEDIIEHY--TYPIILLIDGDKAARKKCYLTQPL 418  
Db 443 ISHLINHLESLSPIV-----SAGSELC-LQOPV 470

```
RESULT 13
US-09-433-353-1
; Sequence 1, Application US/09433353
; Patent No. 6545141
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, TAKESHI
; TITLE OF INVENTION: NOVEL, BRAIN -SPECIFIC ADAPTER MOLECULE, GENE THEREOF,
; AND ANTIBODY THERETO
; FILE REFERENCE: 7898/262241
; CURRENT APPLICATION NUMBER: US/09/433,353
; CURRENT FILING DATE: 1999-11-04
; PRIOR APPLICATION NUMBER: 08/729,416
; PRIOR FILING DATE: 1996-10-11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 474
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-433-353-1

Query Match 7.1%; Score 165.5; DB 2; Length 474;
Best Local Similarity 23.2%; Pred. No. 1.1e-06;
Matches 106; Conservative 60; Mismatches 161; Indels 129; Gaps 25;

QY 13 GFGDLRFQNVSLKNSWPSLSAKGRCRAVLPLPDHRRNLGAVPGGECNNDY--- 69
DB 94 GKSNIQFAGMSISLTISTASLNLRTPDSKQII--ANHMRISISFASGGDP--DTTDYVAY 149
QY 70 --BDP-----EFQLLKAWPSMKILPAPIQSEVADTRYFQDMMEAPLLLPKASVSTERQ 123
DB 150 VAKDPVNRACHILECCDG---LAQDVIGSIGQAFELRFKYLCQCTKIP---ALHDMQ 203
QY 124 TRDVRMTQLEVDKPTFKDVRQRFKGYTKINKTLPPLPPPAITLPPKKYQPLPPAPPE 183
DB 204 SLDEPWTE--EEDG-----GSDHPYNSIPSKMPPPGGFLDTRLKPRPHAP-- 247
QY 184 ESSAYFAPK-----PTPPE-VQRGPRORSKDFSRVLGAEEESHOTKPSSCP 231
DB 248 -DTAQFAGKEQTYQGRHLGDTFGEQWQQTPLRGSSD-----IVSTPEGLHVAAPTGEAP 302
QY 232 SSNQNTQKSPATASSYMPGKHSIQARDHTGSMQHCFAQRCOAASHSPR-----MLPY 286
DB 303 -TYVNTQOIIPP-----QAWPA--AVSSAESSPRKDLFDMKPP 336
QY 287 ENTNSEKP-----DPTKPEK-----DVMQNEWYIGYSRQAVEDVLMKE 326
DB 337 EDALKNQPLGPVLKSAASVEICISPVSPRAPDAKMLELQAEWTYQGEMSKAEGLL--- 393
QY 327 NKDGTFLVRCSTKSKAEP--YVLVVFYGNKVNKIRFLESNQOQFALGTGLRGNEFDS 384
DB 394 EKDGDFLVR---KSTTNPFGSVLTGMHNGQAKHLLLVDP-----GTIRTKDRVDFS 442
QY 385 VEDIIEHY--TYPPILLIDGDKAARRKQCYLTQPL 418
DB 443 ISHLINHLESPLIV-----SAGSELC--LOQPV 470

RESULT 14
US-08-729-416C-7
; Sequence 7, Application US/08729416C
; Patent No. 6013767
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, TAKESHI
; TITLE OF INVENTION: NOVEL BRAIN-SPECIFIC ADAPTER MOLECULE GENE
; AND ANTIBODY THERETO
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PILLSBURY, MADISON & SUTRO, L.L.P.
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA

; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/729,416C
; FILING DATE: 11-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRY, GLENN J.
; REGISTRATION NUMBER: 28458
; REFERENCE/DOCKET NUMBER: 7898/225948
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 594 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-729-416C-7

Query Match 7.1%; Score 165.5; DB 2; Length 594;
Best Local Similarity 23.2%; Pred. No. 1.1e-06;
Matches 106; Conservative 60; Mismatches 161; Indels 129; Gaps 25;

QY 13 GFGDLRFQNVSLKNSWPSLSAKGRCRAVLPLPDHRRNLGAVPGGECNNDY--- 69
DB 214 GKSNIQFAGMSISLTISTASLNLRTPDSKQII--ANHMRISISFASGGDP--DTTDYVAY 269
QY 70 --BDP-----EFQLLKAWPSMKILPAPIQSEVADTRYFQDMMEAPLLLPKASVSTERQ 123
DB 270 VAKDPVNRACHILECCDG---LAQDVIGSIGQAFELRFKYLCQCTKIP---ALHDMQ 323
QY 124 TRDVRMTQLEVDKPTFKDVRQRFKGYTKINKTLPPLPPPAITLPPKKYQPLPPAPPE 183
DB 324 SLDEPWTE--EEDG-----GSDHPYNSIPSKMPPPGGFLDTRLKPRPHAP-- 367
QY 184 ESSAYFAPK-----PTPPE-VQRGPRORSKDFSRVLGAEEESHOTKPSSCP 231
DB 368 -DTAQFAGKEQTYQGRHLGDTFGEQWQQTPLRGSSD-----IVSTPEGLHVAAPTGEAP 422
QY 232 SSNQNTQKSPATASSYMPGKHSIQARDHTGSMQHCFAQRCOAASHSPR-----MLPY 286
DB 423 -TYVNTQOIIPP-----QAWPA--AVSSAESSPRKDLFDMKPP 456
QY 287 ENTNSEKP-----DPTKPEK-----DVMQNEWYIGYSRQAVEDVLMKE 326
DB 457 EDALKNQPLGPVLKSAASVEICISPVSPRAPDAKMLELQAEWTYQGEMSKAEGLL--- 513
QY 327 NKDGTFLVRCSTKSKAEP--YVLVVFYGNKVNKIRFLESNQOQFALGTGLRGNEFDS 384
DB 514 EKDGDFLVR---KSTTNPFGSVLTGMHNGQAKHLLLVDP-----GTIRTKDRVDFS 562
QY 385 VEDIIEHY--TYPPILLIDGDKAARRKQCYLTQPL 418
DB 563 ISHLINHLESPLIV-----SAGSELC--LOQPV 590

RESULT 15
US-09-433-353-7
; Sequence 7, Application US/09433353
; Patent No. 6545141
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, TAKESHI
; TITLE OF INVENTION: NOVEL, BRAIN -SPECIFIC ADAPTER MOLECULE, GENE THEREOF,
; AND ANTIBODY THERETO
; FILE REFERENCE: 7898/262241
; CURRENT APPLICATION NUMBER: US/09/433,353
```

```
; CURRENT FILING DATE: 1999-11-04
; PRIOR APPLICATION NUMBER: 08/729,416
; PRIOR FILING DATE: 1996-10-11
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-433-353-7

Query Match      7.1%; Score 165.5; DB 2; Length 594;
Best Local Similarity 23.2%; Pred. No. 1.8e-06;
Matches 106; Conservative 60; Mismatches 161; Indels 129; Gaps 25;

Qy 13 GFGDLRFQNVSLKRNKSWPSLSAKGRCRAVLEPLDPHRENLAGVPGGKCNNDY--- 69
Db 214 GKNLQFAGMSISLTISTASLNLRTPDSKQII--ANHHMSISFASGGDP--DTTDYVAY 269
Qy 70 --EDP-----FFQLLKAWPSMKILPARPIQSEYADTRYFQDMMEAPLLLPKASVSTERQ 123
Db 270 VAKDPVNRRAACHILECDG---LAQDVIGSIGQAFELRFKQYLQCPTKIP---ALHHRMQ 323
Qy 124 TRDVRMTQLEVDKPTFKVRSQRFKGYTKINKTLPPLPPRAITLPKKYQPLPPAPPE 183
Db 324 SLDEPWTE-BEGD-----GSDHPYNSIPSKMPPGGFLDTRLKPRPHAP-- 367
Qy 184 ESSAYFAPK-----PTPPE-VQGRQRSKDFSRVLGAEEESHOTKPSSSCP 231
Db 368 -DTAQFAGKEQTYVQGRHLGDTFGEDWQQTFLQSGSD---IVSTPEGKLVHAPTGEAP 422
Qy 232 SSNQNTQKSPATASSYMPGKHSIQARDHTGSMQHCPCAPQCAAASHSPR-----MLPY 286
Db 423 -TYVNTQOIIP-----QAWPA--AVSSAESPRKDLFDMKPP 456
Qy 287 ENTNSEKP-----DPTKPEK---DVMQNEWYIGYSRQAVEDVLMKE 326
Db 457 EDALKNQPLGVLKSAASVECSFVSPRAPDAKMLELQAEWTYQGEMSRKEAGLL--- 513
Qy 327 NKDGTFLVRDCSTKSAEP--YVLVVFYGNKYNVNFLESNQFALGTGLRGNEFDS 384
Db 514 EKDGDFLVR---KSTTNGSFVLTMHNGQAKHLLLVDP-----GTIRTKDRVFD 562
Qy 385 VEDIIEHY--TYPEFILLIDGKDAARRKQCYLTOPL 418
Db 563 ISHLINHLESPLIV-----SAGSELC-LOQPV 590
```

Search completed: March 22, 2006, 01:37:32  
Job time : 47 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 22, 2006, 01:36:54 ; Search time 168 Seconds  
(without alignments)  
1081.880 Million cell updates/sec

Title: US-10-717-619-2  
Perfect score: 2316  
Sequence: 1 MTSQGNKRTTKEGFDLRFQ.....QPLPARLLLTQYSSQALHE 435

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA\_Main:  
1: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB.pep:\*  
3: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*  
4: /cgn2\_6/ptodata/1/pubppaa/US10A\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubppaa/US10B\_PUBCOMB.pep:\*  
6: /cgn2\_6/ptodata/1/pubppaa/US11\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2316	100.0	435	5	US-10-717-619-2
2	1316.5	56.8	428	3	US-09-966-955A-4
3	1316.5	56.8	428	4	US-10-104-047-3344
4	1316.5	56.8	443	3	US-09-966-955A-2
5	1306.5	56.4	428	5	US-10-717-619-4
6	1109.5	47.9	353	3	US-09-966-955A-6
7	398	17.2	533	5	US-10-723-860-2611
8	398	17.2	533	5	US-10-287-436A-431
9	398	17.2	533	5	US-10-287-436A-1132
10	319	13.8	456	5	US-10-719-993-671
11	309.5	13.4	433	5	US-10-719-993-672
12	258	11.1	128	5	US-10-617-316-224
13	212	9.2	46	3	US-09-864-761-38652
14	185.5	8.0	503	5	US-10-450-763-53213
15	177	7.6	67	4	US-10-639-067-114
16	161.5	7.0	344	5	US-10-483-506-21
17	159.5	6.9	812	5	US-10-732-923-13470
18	150.5	6.5	812	5	US-10-732-923-13476
19	143	6.2	848	4	US-10-322-281-371
20	143	6.2	881	4	US-10-322-281-374
21	140	6.0	540	5	US-10-723-860-4370
22	138.5	6.0	896	4	US-10-408-765A-1811
23	138.5	6.0	923	4	US-10-258-106-7
24	137	5.9	143	4	US-10-104-047-3542
25	134.5	5.8	715	4	US-10-435-696-98
26	134.5	5.8	3402	5	US-10-626-832-27
27	134	5.8	309	3	US-09-823-187-82

28	133	5.7	410	5	US-10-399-594-15	Sequence 15, Appl
29	133	5.7	675	3	US-09-977-260-4	Sequence 4, Appl
30	133	5.7	675	3	US-09-977-261-4	Sequence 4, Appl
31	133	5.7	675	4	US-10-186-399-3	Sequence 3, Appl
32	133	5.7	675	4	US-10-021-660-108	Sequence 108, App
33	133	5.7	675	4	US-10-211-462-129	Sequence 129, App
34	133	5.7	686	4	US-10-320-801-12	Sequence 12, Appl
35	133	5.7	697	4	US-10-353-690-34	Sequence 34, Appl
36	133	5.7	697	4	US-10-366-288-30	Sequence 30, Appl
37	133	5.7	697	5	US-10-509-599-22	Sequence 22, Appl
38	133	5.7	945	5	US-10-840-060-208	Sequence 208, App
39	133	5.7	945	6	US-11-097-143-25023	Sequence 25023, A
40	132.5	5.7	271	6	US-11-097-143-1605	Sequence 1605, Ap
41	132.5	5.7	592	4	US-10-369-493-1534	Sequence 1534, Ap
42	132.5	5.7	904	6	US-11-097-143-19503	Sequence 19503, A
43	130	5.6	551	4	US-10-369-493-13115	Sequence 13115, A
44	130	5.6	688	4	US-10-081-980B-1	Sequence 1, Appl
45	130	5.6	724	4	US-10-081-980B-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1  
US-10-717-619-2  
; Sequence 2, Application US/10717619  
; Publication No. US20040229237A1  
; GENERAL INFORMATION:  
; APPLICANT: GOITSUKA, RYO  
; TITLE OF INVENTION: A mast cell-specific adapter molecules and cDNAs thereof  
; FILE REFERENCE: 2003-1609/WMC/00653  
; CURRENT APPLICATION NUMBER: US/10717,619  
; CURRENT FILING DATE: 2003-11-21  
; PRIOR APPLICATION NUMBER: PCT/JP00/06351  
; PRIOR FILING DATE: 2000-9-17  
; PRIOR APPLICATION NUMBER: JP11-263778  
; PRIOR FILING DATE: 1999-09-17  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO: 2  
; LENGTH: 435  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-717-619-2

Query Match	100.0%	Score 2316;	DB 5;	Length 435;
Best Local Similarity	100.0%	Pred. No. 7.3e-169;		
Matches	435;	Conservative	0;	Mismatches
		0;	Indels	0;
		Gaps		0;
QY	1	MTSQGNKRTTKEGFDLRFQNVSLKKNRSPWLSAKGRCRAVLEPLPDHRRNLAVPGG	60	
Db	1	MTSQGNKRTTKEGFDLRFQNVSLKKNRSPWLSAKGRCRAVLEPLPDHRRNLAVPGG	60	
QY	61	EKCSNNDYEDPEFQLLKAWPSMKIIPARIQSEYADTRYFQDMMEAPLLPPKASVST	120	
Db	61	EKCSNNDYEDPEFQLLKAWPSMKIIPARIQSEYADTRYFQDMMEAPLLPPKASVST	120	
QY	121	EROTRDRVMTQLEVDKPTFKDVSQRFKGYTKTKNTPLPPRPAITLPPKYQPLPPA	180	
Db	121	EROTRDRVMTQLEVDKPTFKDVSQRFKGYTKTKNTPLPPRPAITLPPKYQPLPPA	180	
QY	181	PPESAYFAPKPTTFPEVQGRQPSAKDFSRVLGAEEESHHTQKPESSCPSSNQNTOKS	240	
Db	181	PPESAYFAPKPTTFPEVQGRQPSAKDFSRVLGAEEESHHTQKPESSCPSSNQNTOKS	240	
QY	241	PPAIASSYMPGKHSIQARDHTGSMOHCPCQAORCAAAASHSPRMLPYENTNSEKPDTPKPD	300	
Db	241	PPAIASSYMPGKHSIQARDHTGSMOHCPCQAORCAAAASHSPRMLPYENTNSEKPDTPKPD	300	
QY	301	EKDVQNEWYIGSYRQAVEDVLKMKENKDGTFVLVDCSTKSKABPYVVLVYFGNKVNVK	360	
Db	301	EKDVQNEWYIGSYRQAVEDVLKMKENKDGTFVLVDCSTKSKABPYVVLVYFGNKVNVK	360	

```

Qy 361 IRFLESNQFALGTGLRGNEFDSVEDIIHYHTYFPILLIDGDKAARRKQCYLTQPL 420
Db 361 IRFLESNQFALGTGLRGNEFDSVEDIIHYHTYFPILLIDGDKAARRKQCYLTQPL 420

Qy 421 ARLLLTQYSSQALHE 435
Db 421 ARLLLTQYSSQALHE 435

RESULT 2
US-09-966-955A-4
; Sequence 4, Application US/09966955A
; Patent No. US20020155563A1
; GENERAL INFORMATION:
; APPLICANT: Perez-Villar, Juan J.
; APPLICANT: Chang, Han
; APPLICANT: Yang, Wen-Pin
; APPLICANT: Wu, Yuli
; APPLICANT: Whitney, Gena S.
; APPLICANT: Kanner, Steven B.
; TITLE OF INVENTION: Identification and Cloning of a Full-length Human
; TITLE OF INVENTION: Clink-related Gene, MIST (Mast Cell Immunoreceptor
; TITLE OF INVENTION: Signal Transducer)
; FILE REFERENCE: 3053-4113US1
; CURRENT APPLICATION NUMBER: US/09/966,955A
; CURRENT FILING DATE: 2001-09-28
; PRIOR FILING DATE: 2001-09-28
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 428
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; OTHER INFORMATION: HUMAN MIST SPLICED VARIANT CLONE #7, AMINO ACID
; OTHER INFORMATION: SEQUENCE
US-09-966-955A-4

```

```

Query Match 56.8%; Score 1316.5; DB 3; Length 428;
Best Local Similarity 62.1%; Pred. No. 2.6e-92;
Matches 266; Conservative 46; Mismatches 111; Indels 5; Gaps 4;

Qy 1 MTSQGNKRTTKEGDLRFQNVSLKRWSPSLSSAKGRCAVLEPLPDHRRNLAVPGG 60
Db 1 MNRQGNKRTTKEGSDNLKFNFLPKNRSWPRINSATGQYQRMNKPLLDWERNFAAVLDG 60

Qy 61 EKCSNNDYEDPFPQLLKAWPSMKILPARPIQSEYADTRYFODMMEAPLLLPKASVST 120
Db 61 AKGSHDDDDYDDELRLMEETWQSIKILPARPIKSEYADTRYFVKVAMDTPPLDTRTSISI 120

Qy 121 ERQTRDVRMTQLEBEVDKPTFKDVRQSRQKFKYTKINKTLPPLPPPAITLPPKYQPLPPA 180
Db 121 GQPTWNTQ-TRLERVDKPIKSDVRQSRQKFKYTKINKTLPPLPPPAITLPPKYQPLPP- 178

Qy 181 PPESSAYFAPKPTFPFVQGRQSRQKFKYTKINKTLPPLPPPAITLPPKYQPLPPA 240
Db 179 EPSSRPPLSQRTTFPEVQRMPSQISLRDLSEVLEAEKVPNQRKPESTHLENQNTQEI 238

Qy 241 PPAIASSSYMPGKHISQARDHTGSMQCHPAQRQAAASHSP--RMLPYENTNSEKDPDK 298
Db 239 PLAISSSFTTNSHVSQNRDRHGMQPCSPQRCQPPASCSPHENILPKYKTSWRPFPK 298

Qy 299 PDEKDVQNEWYIGEYSRQAVEAFMKNKDGFTFLVRDCSTKSKAEPYVLVVFYGNKVYN 358
Db 299 SDRKDVQNEWYIGEYSRQAVEAFMKNKDGFTFLVRDCSTKSKAEPYVLVVFYGNKVYN 358

Qy 359 VKIRFLESNQFALGTGLRGNEFDSVEDIIHYHTYFPILLIDGDK-AARRKQCYLTOP 417
Db 359 VKIRFLERNQFALGTGLRGDEKFDSDIEDIIHYKNFPILLIDGDKKTGVHRKQCHLTOP 418

Qy 418 LPLRLILL 425
Db 419 LPLTRHLL 426

```

```

RESULT 4
US-09-966-955A-2
; Sequence 2, Application US/09966955A
; Patent No. US20020155563A1
; GENERAL INFORMATION:
; APPLICANT: Perez-Villar, Juan J.
; APPLICANT: Chang, Han
; APPLICANT: Yang, Wen-Pin
; APPLICANT: Wu, Yuli
; APPLICANT: Whitney, Gena S.
; APPLICANT: Kanner, Steven B.
; TITLE OF INVENTION: Identification and Cloning of a Full-length Human
; TITLE OF INVENTION: Clink-related Gene, MIST (Mast Cell Immunoreceptor

```

```

Db 419 LPLTRHLL 426

RESULT 3
US-10-104-047-3344
; Sequence 3344, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1e1 full length cdna
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3344
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-104-047-3344

```

```

Query Match 56.8%; Score 1316.5; DB 4; Length 428;
Best Local Similarity 62.1%; Pred. No. 2.6e-92;
Matches 266; Conservative 46; Mismatches 111; Indels 5; Gaps 4;

Qy 1 MTSQGNKRTTKEGDLRFQNVSLKRWSPSLSSAKGRCAVLEPLPDHRRNLAVPGG 60
Db 1 MNRQGNKRTTKEGSDNLKFNFLPKNRSWPRINSATGQYQRMNKPLLDWERNFAAVLDG 60

Qy 61 EKCSNNDYEDPFPQLLKAWPSMKILPARPIQSEYADTRYFODMMEAPLLLPKASVST 120
Db 61 AKGSHDDDDYDDELRLMEETWQSIKILPARPIKSEYADTRYFVKVAMDTPPLDTRTSISI 120

Qy 121 ERQTRDVRMTQLEBEVDKPTFKDVRQSRQKFKYTKINKTLPPLPPPAITLPPKYQPLPPA 180
Db 121 GQPTWNTQ-TRLERVDKPIKSDVRQSRQKFKYTKINKTLPPLPPPAITLPPKYQPLPP- 178

Qy 181 PPESSAYFAPKPTFPFVQGRQSRQKFKYTKINKTLPPLPPPAITLPPKYQPLPPA 240
Db 179 EPSSRPPLSQRTTFPEVQRMPSQISLRDLSEVLEAEKVPNQRKPESTHLENQNTQEI 238

Qy 241 PPAIASSSYMPGKHISQARDHTGSMQCHPAQRQAAASHSP--RMLPYENTNSEKDPDK 298
Db 239 PLAISSSFTTNSHVSQNRDRHGMQPCSPQRCQPPASCSPHENILPKYKTSWRPFPK 298

Qy 299 PDEKDVQNEWYIGEYSRQAVEAFMKNKDGFTFLVRDCSTKSKAEPYVLVVFYGNKVYN 358
Db 299 SDRKDVQNEWYIGEYSRQAVEAFMKNKDGFTFLVRDCSTKSKAEPYVLVVFYGNKVYN 358

Qy 359 VKIRFLESNQFALGTGLRGNEFDSVEDIIHYHTYFPILLIDGDK-AARRKQCYLTOP 417
Db 359 VKIRFLERNQFALGTGLRGDEKFDSDIEDIIHYKNFPILLIDGDKKTGVHRKQCHLTOP 418

Qy 418 LPLRLILL 425
Db 419 LPLTRHLL 426

```

```
; TITLE OF INVENTION: Signal Transducer)
; FILE REFERENCE: 3053-4113US1
; CURRENT APPLICATION NUMBER: US/09/966,955A
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/237030
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 443
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; OTHER INFORMATION: HUMAN FULL-LENGTH MIST cDNA CLONE #8, TRANSLATED
; OTHER INFORMATION: AMINO ACID SEQUENCE
US-09-966-955A-2

Query Match          56.8%; Score 1316.5; DB 3; Length 443;
Best Local Similarity 62.1%; Pred. No. 2.8e-92;
Matches 266; Conservative 46; Mismatches 111; Indels 5; Gaps 4;

Qy 1 MTSQGNKRTTKEGFDLRFQNVSLLNKRSWPSLSAKGRCAVLEPLPDHRRNLGAVPGG 60
Db 16 MNRQGNKRTTKEGNDLKQNFSLPKNRSWPRINSATGQYQRMNKPLLDWERNFAA 75
Qy 61 EKNSNNDDYEDPEFOLLKAWPSMKILPARPIQSEYADTRYFQDMMEAPLLPPKASVST 120
Db 76 AKGSHDDDDYDDPELMEETWQSIKILPARPIKSEYADTRYFKVAMDTPLDTRTSISI 135
Qy 121 ERQTRDVRMTQLEEVDPKTFKVSQRQFSGKFKYTKINKTLPPLPPPAITLPPKYQPLPPA 180
Db 136 GQPTWNTQ-TRLEKVDKPIKSRQVNIKGDASVRKNKIPLPPRPLITLPPKYQPLPP- 193
Qy 181 PPESSAYFAPKPTFFEVQGRPRQSAKDFSRVLGAEEESHQTKPSSCPSSNQNTQKS 240
Db 194 EPSSRRPPLSQRHTFFEVQMPQSI SLRDLSEVLEAEKVPNHQKPESTHLENQNTQEI 253
Qy 241 PPAIASSSYMPGKHISIQARDHTGSMOHCAPQRCQAASHSP--RMLPYENTNSEKPDPTK 298
Db 254 PLAISSSFTTSNHSVQNRDHRGMPQSPQRCQPPASCSPHENILPYKTSWRPPPKR 313
Qy 299 PDEKDVQWNEWYIGEYSRQAVEDVLMEKNKDGTFVLVDCSTKSKAEPPYVLVVFYGNKVYN 358
Db 314 SDRKDVQWNEWYIGEYSRQAVEAEAFMKNKDGSLVDCSTKSKBEPYVLAVFYENKVYN 373
Qy 359 VKIRFLSNQOFPALGTGLRGNEMFDSVEDIIHYHTYTPILLIDGKDK-AARRKQCYLTOP 417
Db 374 VKIRFLERNQOFPALGTGLRGEKFDSDVEDIIHYKNFPILLIDGKDKTGVRHKKOCHLTQP 433
Qy 418 LPLARLLL 425
Db 434 LPLTRHLL 441

RESULT 5
US-10-717-619-4
; Sequence 4, Application US/10717619
; Publication No. US20040229237A1
; GENERAL INFORMATION:
; APPLICANT: GOITSUKA, Ryo
; TITLE OF INVENTION: A mast cell-specific adapter molecules and cDNAs thereof
; FILE REFERENCE: 2003-1609/WMC/00653
; CURRENT APPLICATION NUMBER: US/10/717,619
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: PCT/JP00/06351
; PRIOR FILING DATE: 2000-9-17
; PRIOR APPLICATION NUMBER: JP11-263778
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 428
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
US-10-717-619-4

Query Match          56.4%; Score 1306.5; DB 5; Length 428;
Best Local Similarity 61.7%; Pred. No. 1.5e-91;
Matches 264; Conservative 47; Mismatches 112; Indels 5; Gaps 4;

Qy 1 MTSQGNKRTTKEGFDLRFQNVSLLNKRSWPSLSAKGRCAVLEPLPDHRRNLGAVPGG 60
Db 1 MNRQGNKRTTKEGNDLKQNFSLPKNRSWPRINSATGQYQRMNKPLLDWERNFAA 60
Qy 61 EKNSNNDDYEDPEFOLLKAWPSMKILPARPIQSEYADTRYFQDMMEAPLLPPKASVST 120
Db 61 AKGSHDDDDYDDPELMEETWQSIKILPARPIKSEYADTRYFKVAMDTPLDTRTSISI 120
Qy 121 ERQTRDVRMTQLEEVDPKTFKVSQRQFSGKFKYTKINKTLPPLPPPAITLPPKYQPLPPA 180
Db 121 GQPTWNTQ-TRLEKVDKPIKSRQVNIKGDASVRKNKIPLPPRPLITLPPKYQPLPP- 178
Qy 181 PPESSAYFAPKPTFFEVQGRPRQSAKDFSRVLGAEEESHQTKPSSCPSSNQNTQKS 240
Db 179 EPSSRRPPLSQRHTFFEVQMPQSI SLRDLSEVLEAEKVPNHQKPESTHLENQNTQEI 238
Qy 241 PPAIASSSYMPGKHISIQARDHTGSMOHCAPQRCQAASHSP--RMLPYENTNSEKPDPTK 298
Db 239 PLAISSSFTTSNHSVQNRDHRGMPQSPQRCQPPASCSPHENILPYKTSWRPPPKR 298
Qy 299 PDEKDVQWNEWYIGEYSRQAVEDVLMEKNKDGTFVLVDCSTKSKAEPPYVLVVFYGNKVYN 358
Db 299 SDRKDVQWNEWYIGEYSRQAVEAEAFMKNKDGSLVDCSTKSKBEPYVLAVFYENKVYN 358
Qy 359 VKIRFLSNQOFPALGTGLRGNEMFDSVEDIIHYHTYTPILLIDGKDK-AARRKQCYLTOP 417
Db 359 VKIRFLERNQOFPALGTGLRGEKFDSDVEDIIHYKNFPILLIDGKDKTGVRHKKOCHLTQP 418
Qy 418 LPLARLLL 425
Db 419 LPLTRHLL 426

RESULT 6
US-09-966-955A-6
; Sequence 6, Application US/09966955A
; Patent No. US20020155563A1
; GENERAL INFORMATION:
; APPLICANT: Perez-Villar, Juan J.
; APPLICANT: Chang, Han
; APPLICANT: Yang, Wen-Pin
; APPLICANT: Wu, Yuli
; APPLICANT: Whitney, Gena S.
; APPLICANT: Kanner, Steven B.
; TITLE OF INVENTION: Identification and Cloning of a Full-length Human
; TITLE OF INVENTION: Clnk-related Gene, MIST (Mast Cell Immunoreceptor
; TITLE OF INVENTION: Signal Transducer)
; FILE REFERENCE: 3053-4113US1
; CURRENT APPLICATION NUMBER: US/09/966,955A
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/237030
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 6
; LENGTH: 353
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; OTHER INFORMATION: HUMAN MIST SPLICE VARIANT CLONE #12, TRANSLATED
; OTHER INFORMATION: AMINO ACID SEQUENCE
US-09-966-955A-6

Query Match          47.9%; Score 1109.5; DB 3; Length 353;
Best Local Similarity 64.4%; Pred. No. 1.5e-76;
Matches 226; Conservative 31; Mismatches 89; Indels 5; Gaps 4;
```





```
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US 60/336,220
; PRIOR FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 1446
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 1132
; LENGTH: 533
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-287-436A-1132

Query Match      17.2%; Score 398; DB 5; Length 533;
Best Local Similarity 30.2%; Pred. No. 8e-22;
Matches 124; Conservative 52; Mismatches 146; Indels 88; Gaps 14;

Qy 68 DYEDPEFQLKAWPSMKILPARPIQSE--YADTRYFQDMMEAP-----LLPPKA 116
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 144 DYRPPPSNDEALQN-SILPAKPPNSMYIDRPPSGKTPQQPVPPQRPMAALPPPA 202
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 117 -----SVSTERQTRDVRMTQLEVD-----KPT 139
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 203 GRNHSPLPPQTWHEEPSRSRHHKAKLPAPSIDRSTKPLDRSLAPFDREFTLGKKPP 262
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 140 FKD---VRSORFGKFKYTKNKTLPLPPRAITLPPKYQPLPPAPPESSAYFAPKPTFP 196
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 263 FSKPSPISAGRSJGEHLPKIQKPLPP-----TTERHERSSDLP-----GKPPVP 308
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 197 EVQGRQRSQAKDSFSLVGAEEESHQTKPE-----SSCFSSNQNTQKSPPAIASSSYM 250
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 309 KHGSGPDRRE-----NDEDDVHORPLQPALPMSNTFFSRSTKPSPMNPLPSSH 360
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 251 PGKHSIQARDHTGSMQHCPCQACQAAASHSP-----RMLPYENTNSEKPDPTPKDEKDVW 305
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 361 PGAFS-BSNSFFQASASLPYFSGQPSNRPPRAEGRNFPLPLPNKPRP-PSPAEEENSL 418
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 306 QNEWYIGEYSROAVEDVLKENKDGTFELVDCSTKSKABPYLVVFGNKNVNVKIRFLE 365
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 419 NEEMVYSYITRPAEALRKINQDGTFLVDRSSKKTNTNPNYLVMLVKDKVYNIQIRYQK 478
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 366 SNQOQFALGTGLRGNEFDSVEDIEHYTYFPILLIDGDKAARRKQCYLT 415
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 479 ESQVYLLGTGLRGKEDFLSVSDIIDYERKMPPLLLIDGKNRGS-RYQCTLT 527
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

RESULT 10
US-10-719-993-671
; Sequence 671, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001496
; CURRENT APPLICATION NUMBER: US/10/719,993
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 55342
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 671
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-719-993-671

Query Match      13.8%; Score 319; DB 5; Length 456;
Best Local Similarity 25.0%; Pred. No. 7.4e-16;
Matches 99; Conservative 58; Mismatches 139; Indels 100; Gaps 12;

Qy 64 NSNNYDEDPFQLKAWPSMKILPARPIQSEYADTRYFQDMMEAP---LLLPPKASVST 120
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 91 NADDSYEPPEV-----QETRPVHPALPFARGEYIDNRSSQ--RHSPPFSKTLSPKSPWS 144
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 121 ERQTRDVRMTQLEVDKPTFKQVRSORFKGYTKINKTLPPLPPRAITLPPKYQPLPPA 180
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 145 E-----KARLTSTLPALTALQKPQ-VPPK 167
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 181 PP---BESSAYFAPKPTFFEVQGRQRSQAKDSFSLVGAEEESHQTKPESSCPSSNQ- 236
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 168 PKGLLEADYVVP-----VEDNDENYIHTPESSPPPEKGRNS 206
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 237 ---TQSPPAIASSSYMPGKH-----SIQARDHTGSM---QHCPCQACQAA----- 276
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 207 GAWETKSPPPAAPSPPLPRAGKKTTPPLKTPPVASQQNASSVCEKEKPIPAERHGRSS 266
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 277 ASHSRMLPYENTNSEKPD-----TKPDEKDVWQNEWYIG 312
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 267 AVQSPVPFPAQKQIHQKPIPLPRFTTEGNGNPTVDGPLPSFSSNSTISQEAQVLCPPYAG 326
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 313 EYSROQAVEDVLKENKDGTFELVDCSTKSKABPYLVVFGNKNVNVKIRFLESNQFAL 372
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 327 ACDKSGAEALHRSNKGSLIRKSGHDSKQPYTLVVFNFKNVYINIPVRFIEATKQYAL 386
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 373 GTGLRGNEFDSVEDIEHYTYFPILLIDGK 404
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 387 GRKNKGEYFGSVAEIRNHQHSPLVLIDSQN 418
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
```

```
Db 145 EKARLTSTLPALTALQKPQV-----PPKPKGLLEADYVVPV 182
Qy 181 PPESSAYFAPKPTFFEVQGRQRSQAKDSFSLVGAEEESHQTKPESSCPSSNQNT--- 237
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 183 EDNDENYIHTPESSSPPEKAPM-----VNRSTKPNSSTPASPPGTASG 226
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 238 -----QKSPPAIASSYMP--GKH-----SIQARDHTGSM---QHCPCQACQAA- 276
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 227 RNSGAWETKSPPP-AAPSPPLPRAGKKTTPPLKTPPVASQQNASSVCEKEKPIPAERHGRSS 285
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 277 ---ASHSRMLPYENTNSEKPD-----TKPDEKDVWQNE 308
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 286 HRQEAQVQSPVPFPAQKQIHQKPIPLPRFTTEGNGNPTVDGPLPSFSSNSTISQEAQVLCKP 345
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 309 WYIGEYSROQAVEDVLKENKDGTFELVDCSTKSKABPYLVVFGNKNVNVKIRFLESNQ 368
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 346 WYAGACDRKSAEALHRSNKGSLIRKSGHDSKQPYTLVVFNFKNVYINIPVRFIEATK 405
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 369 QFALGTGLRGNEFDSVEDIEHYTYFPILLIDGK 404
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 406 QYALGRKNKGEYFGSVAEIRNHQHSPLVLIDSQN 441
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

RESULT 11
US-10-719-993-672
; Sequence 672, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001496
; CURRENT APPLICATION NUMBER: US/10/719,993
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 55342
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 672
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-719-993-672

Query Match      13.4%; Score 309.5; DB 5; Length 433;
Best Local Similarity 25.3%; Pred. No. 3.7e-15;
Matches 99; Conservative 51; Mismatches 127; Indels 115; Gaps 12;

Qy 64 NSNNYDEDPFQLKAWPSMKILPARPIQSEYADTRYFQDMMEAP---LLLPPKASVST 120
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 91 NADDSYEPPEV-----QETRPVHPALPFARGEYIDNRSSQ--RHSPPFSKTLSPKSPWS 144
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 121 ERQTRDVRMTQLEVDKPTFKQVRSORFKGYTKINKTLPPLPPRAITLPPKYQPLPPA 180
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 145 E-----KARLTSTLPALTALQKPQ-VPPK 167
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 181 PP---BESSAYFAPKPTFFEVQGRQRSQAKDSFSLVGAEEESHQTKPESSCPSSNQ- 236
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 168 PKGLLEADYVVP-----VEDNDENYIHTPESSPPPEKGRNS 206
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 237 ---TQSPPAIASSSYMPGKH-----SIQARDHTGSM---QHCPCQACQAA----- 276
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 207 GAWETKSPPPAAPSPPLPRAGKKTTPPLKTPPVASQQNASSVCEKEKPIPAERHGRSS 266
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 277 ASHSRMLPYENTNSEKPD-----TKPDEKDVWQNEWYIG 312
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 267 AVQSPVPFPAQKQIHQKPIPLPRFTTEGNGNPTVDGPLPSFSSNSTISQEAQVLCPPYAG 326
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 313 EYSROQAVEDVLKENKDGTFELVDCSTKSKABPYLVVFGNKNVNVKIRFLESNQFAL 372
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 327 ACDKSGAEALHRSNKGSLIRKSGHDSKQPYTLVVFNFKNVYINIPVRFIEATKQYAL 386
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 373 GTGLRGNEFDSVEDIEHYTYFPILLIDGK 404
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 387 GRKNKGEYFGSVAEIRNHQHSPLVLIDSQN 418
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
```

## RESULT 12

US-10-617-316-224  
; Sequence 224, Application US/10617316  
; Publication No. US20050130263A1  
; GENERAL INFORMATION:  
; APPLICANT: Havukkala, Ilkka J  
; APPLICANT: Glenn, Matthew  
; APPLICANT: Grigor, Murray R.  
; APPLICANT: Molenaar, Adrian J.  
; TITLE OF INVENTION: Compositions Isolated From Bovine  
; TITLE OF INVENTION: Mammary Gland and Methods For Their Use.  
; FILE REFERENCE: 11000.1046UIC1  
; CURRENT APPLICATION NUMBER: US/10/617,316  
; CURRENT FILING DATE: 2003-07-09  
; PRIOR APPLICATION NUMBER: US 09/699,146  
; PRIOR FILING DATE: 2000-10-27  
; PRIOR APPLICATION NUMBER: US 60/162,701  
; PRIOR FILING DATE: 1999-10-29  
; NUMBER OF SEQ ID NOS: 262  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 224  
; LENGTH: 128  
; TYPE: PRT  
; ORGANISM: Bovine  
US-10-617-316-224

Query Match 11.1%; Score 258; DB 5; Length 128;  
Best Local Similarity 42.9%; Pred. No. 6.6e-12;  
Matches 54; Conservative 21; Mismatches 45; Indels 6; Gaps 1;

Qy 272 RCQAAASHSPRLPYENTNSEKDDPTKP-----DEKDVWQNEWYIGEYSRQAVEDVLMK 325  
Db 3 KCYTFKPGNRLPRMEGRNFPSPKPSQSGEEENSLNEDWYVSVYVTRTEAAALRK 62  
Qy 326 ENKDGTEFLVDCSTKSKAEPPVVLVVFYGNKVYVVKIRFLESNQOQFALGTGLRGNEPDSV 395  
Db 63 INQDGTFLVDRSKKTSINPVLVYKDKVYNIQIRYQESQVYLLGTGLRGKEDFLSV 122

Qy 386 EDIIEH 391  
Db 123 SDIIDY 128

## RESULT 13

US-09-864-761-38652  
; Sequence 38652, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; \*TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aeomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 38652  
; LENGTH: 46  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC005599.5  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.4  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.3  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.6  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.9  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.4  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 5.7  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.3  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.4  
; OTHER INFORMATION: SWISSPROT HIT: Q60787, EVALU 9.00e-10  
US-09-864-761-38652

Query Match 9.2%; Score 212; DB 3; Length 46;  
Best Local Similarity 89.1%; Pred. No. 5.8e-09;  
Matches 41; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 329 DGTFLVDRDCSTKSKAEPPVVLVVFYGNKVYVVKIRFLESNQOQFALGT 374  
Db 1 DGSFLVDRDCSTKSKAEPPVVLVVFYGNKVYVVKIRFLESNQOQFALGT 46

## RESULT 14

US-10-450-763-53213  
; Sequence 53213, Application US/10450763  
; Publication No. US20050196754A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 790CIP3/US  
; CURRENT APPLICATION NUMBER: US/10/450,763  
; CURRENT FILING DATE: 2003-06-11  
; PRIOR APPLICATION NUMBER: PCT/US01/08631  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 09/540,217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/649,167  
; PRIOR FILING DATE: 2000-08-23  
; NUMBER OF SEQ ID NOS: 60736  
; SOFTWARE: Custom  
; SEQ ID NO 53213  
; LENGTH: 503  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:



**This Page Blank (Uplo)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 22, 2006, 01:37:44 ; Search time 23 Seconds  
(without alignments)  
541.346 Million cell updates/sec

Title: US-10-717-619-2

Perfect score: 2316

Sequence: 1 MTSQGNKRTTKGFGDLRFQ.....QPLPLARLLLTQYSSQALH 435

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 169630 seqs, 28622889 residues

Total number of hits satisfying chosen parameters: 169630

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: 1: Published Applications AA New.\*

2: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1316.5	56.8	428	7	US-11-072-512-3344
2	138.5	6.0	950	6	US-10-501-035-357
3	137	5.9	143	7	US-11-072-512-3542
4	136.5	5.9	561	6	US-10-467-033-2
5	136.5	5.9	561	6	US-10-467-033-6
6	128.5	5.5	393	7	US-11-221-263-26
7	128	5.5	724	7	US-11-109-156-19
8	127.5	5.5	863	7	US-11-169-041-167
9	121.5	5.2	2344	6	US-10-330-773-627
10	117	5.1	428	7	US-11-221-263-24
11	116.5	5.0	536	6	US-10-821-231C-1
12	116	5.0	667	6	US-10-821-234-1477
13	116	5.0	667	7	US-11-124-368A-250
14	116	5.0	686	7	US-11-124-368A-239
15	116	5.0	708	7	US-11-072-175-158
16	116	5.0	726	7	US-11-124-368A-247
17	116	5.0	722	7	US-11-124-368A-248
18	116	5.0	750	7	US-11-124-368A-244
19	116	5.0	750	7	US-11-124-368A-249
20	116	5.0	777	7	US-11-124-368A-238
21	116	5.0	778	7	US-11-124-368A-240
22	116	5.0	791	7	US-11-124-368A-245
23	115.5	5.0	565	6	US-10-915-161-2
24	115	5.0	2801	6	US-10-330-773-630
25	114.5	4.9	533	7	US-11-230-995-3

26	114	4.9	460	7	US-11-087-099-9781	Sequence 9781, Ap
27	112.5	4.9	798	6	US-10-514-531-2	Sequence 2, Appli
28	112.5	4.9	941	7	US-11-169-232-14	Sequence 14, Appl
29	112.5	4.9	1022	7	US-11-169-232-84	Sequence 84, Appl
30	112.5	4.9	1038	7	US-11-169-232-74	Sequence 74, Appl
31	112.5	4.9	1049	7	US-11-169-232-58	Sequence 58, Appl
32	112.5	4.9	1140	7	US-11-169-232-104	Sequence 104, App
33	112.5	4.9	1270	7	US-11-169-232-44	Sequence 44, Appl
34	112.5	4.9	1311	7	US-11-169-232-42	Sequence 42, Appl
35	112.5	4.9	1313	7	US-11-169-232-142	Sequence 142, App
36	112.5	4.9	1314	7	US-11-169-232-50	Sequence 50, Appl
37	112.5	4.9	1320	7	US-11-169-232-46	Sequence 46, Appl
38	112.5	4.9	1320	7	US-11-169-232-60	Sequence 60, Appl
39	112.5	4.9	1354	7	US-11-169-232-48	Sequence 48, Appl
40	112.5	4.9	1361	7	US-11-169-232-40	Sequence 40, Appl
41	112.5	4.9	1363	7	US-11-169-232-52	Sequence 52, Appl
42	112.5	4.9	1404	7	US-11-169-232-2	Sequence 2, Appli
43	112.5	4.9	1404	7	US-11-169-232-62	Sequence 62, Appl
44	111.5	4.8	410	7	US-11-096-568A-22820	Sequence 22820, A
45	111.5	4.8	428	7	US-11-096-568A-22819	Sequence 22819, A

#### ALIGNMENTS

#### RESULT 1

US-11-072-512-3344  
; Sequence 3344, Application US/11072512  
; Publication No. US20060029945A1

#### GENERAL INFORMATION:

; APPLICANT: ISOGAI, TAKAO  
; APPLICANT: SUGIYAMA, TOMOYASU  
; APPLICANT: OTSUKI, TETSUJI  
; APPLICANT: WAKAMATSU, AI  
; APPLICANT: SATO, HIROYUKI  
; APPLICANT: ISHII, SHIZUKO  
; APPLICANT: YAMAMOTO, JUN-ICHI  
; APPLICANT: ISONO, YUUKO  
; APPLICANT: HIO, YURI  
; APPLICANT: OTSUKA, KAORU  
; APPLICANT: NAGAI, KEIICHI  
; APPLICANT: IRIE, RYOTARO  
; APPLICANT: TAMECHIKA, ICHIRO  
; APPLICANT: SEKI, NAOHICO  
; APPLICANT: YOSHIKAWA, TSUTOMU  
; APPLICANT: OTSUKA, MOTOYUKI  
; APPLICANT: NAGAHARI, KENJI  
; APPLICANT: MASUHO, YASUHIKO  
; TITLE OF INVENTION: Novel full length cDNA  
; FILE REFERENCE: 084335-0191  
; CURRENT APPLICATION NUMBER: US/11/072,512  
; CURRENT FILING DATE: 2005-03-07  
; PRIOR APPLICATION NUMBER: US 60/350,978  
; PRIOR FILING DATE: 2002-01-25  
; PRIOR APPLICATION NUMBER: JP 2001-379298  
; PRIOR FILING DATE: 2001-11-05  
; NUMBER OF SEQ ID NOS: 4096  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3344  
; LENGTH: 428  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-072-512-3344

Query Match 56.8%; Score 1316.5; DB 7; Length 428;

Best Local Similarity 62.1%; Pred. No. 2.8e-100;

Matches 266; Conservative 46; Mismatches 111; Indels 5; Gaps 4;

Qy 1 MTSQGNKRTTKGFGDLRFQNSVLLKNRSPSSAKRCRAVLPLPDHRRNLGYPGG 60

Db 1 MNEQGNKRTTKGSGNDKLFQNFSLPKNRSWPINSATQYQRMNKPDLDERNFAAVLDG 60

Qy 61 EKCNSNDYEDPDEPQLLKAWPSMKILLPARPIQSEYADTRYFQDMMEAPLLLPKASVST 120

```
Db 61 AKGSDDDYDDPELRMEETWQSIKILPARPIKESYADTHYFKVAMDTPLPLTRTSISI 120
QY 121 ERQTRDVRMTQLEVDKPTFKDVRQRKFGKTKINKTLPPLPPRAITLPPKKYQPLPPA 180
Db 121 GQPTWNTQ-TRLERVDKPIKSDVRQNSQIKGDASVRKNKIPLPPRPPLITLPPKKYQPLPP- 178
QY 181 PPESSAYFAPKPTFPVQGRQRORSKDFSRVLGAEESHQTKPESSCPSSNONTOKS 240
Db 179 EPSSRPPLSRHTFPVQRMPSQISLRDISELEAKVPHNQKPESTHLLNQNQTQEI 238
QY 241 PPAIASSYMPGKHISQIARDHTSMOHCPCQORQAAASHSP--RMLPYENTNSEKDPDK 298
Db 239 PLAISSSFTTSHSVQNRDHRGMQPCSPQRCQPPASCSPHENILPYKTSWRPFPKR 298
QY 299 PDEKDVQNWQEWYIGEYSRQAVEDVLMKENKDGTFVLVRDCSTKSKABPYVLVVFYGNKVYN 358
Db 299 SDRKDVQHNWYIGEYSRQAVEAFMKNKDGSLVRDCSTKSKBEPYVLAVFVFNKVYN 358
QY 359 VKIRFLESNOQFALGTGLRGNEMFDSVEDIIEHYTFPILLIDGDK-AARRKQCYLTOP 417
Db 359 VKIRFLENOQFALGTGLRDEKFDSEDIIEHYKNFPILLIDGDKGTGVHRKQCHLTOP 418
QY 418 LPLARLLL 425
Db 419 LPLTRHLL 426

RESULT 2
US-10-501-035-357
; Sequence 357, Application US/10501035
; Publication No. US20060046249A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES AND POLYPEPTIDE FOR PREDICTING
; TITLE OF INVENTION: ACTIVITY OF COMPOUNDS THAT INTERACT WITH PROTEIN TYROSINE KINASE
; TITLE OF INVENTION: AND/OR PROTEIN TYROSINE KINASE PATHWAYS
; FILE REFERENCE: D0185 PCT
; CURRENT APPLICATION NUMBER: US/10/501,035
; CURRENT FILING DATE: 2004-07-09
; PRIOR APPLICATION NUMBER: US 60/350,061
; PRIOR FILING DATE: 2002-01-18
; NUMBER OF SEQ ID NOS: 795
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 357
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-501-035-357

Query Match 6.0%; Score 138.5; DB 6; Length 950;
Best Local Similarity 22.8%; Pred. No. 0.0023;
Matches 80; Conservative 41; Mismatches 153; Indels 77; Gaps 15;

QY 50 HRENLAGVPG-----GKCNNDYEDPEFO-LLKAMP-----SMKI 85
Db 625 HLQNLARTKGILELNKVLLEYIGPEADPNLAPAPQLQDLASSCPQEVSVSQQESVST 684
QY 86 LPAR---PIQESYADTRYFQDMWAEPLLL-----PPKASVSTERQTRDVRM 129
Db 685 LPASVHPQLSPROSLETQYLQHRLOKPELSLSKAQNTCOLYCKEPPR---SLEQQLQEHRL 741
QY 130 TQLEEVDPKTFKDVQRGKFGYKINKTLPPLPPRAITLPPKKYQPLPPAPPESSAYF 189
Db 742 QQ-----KRLFLQKSOLOAYFNQMQIAESSYPQSQQLPLPR--QETPP-PSQAPPFS 793
QY 190 AKPTPEVQGRQRORSKDFSRVLGAEESHQTKPESS-----CPSSNONTOKSP 242
Db 794 LTQPLSVLEPSSSEQ---MQYSFPLSQYQEMQIQLPSTSGPRAAPLPTLQLOQQQPPPP 850
QY 243 AIASSSYMPGKHISQIARDHTGSMQHCPCQORQAAASHSPRMLPYENTNSEKDPDK 302
Db 851 PPPPPRQPCGAAPLQ---FSYQTCELPSAASPAPDYPPTCOYPVDGAQQSLLTGPD-- 905
```

```
QY 303 DVQWQNEWYIGEYSRQAVEDVLMKENKDGTFVLVRDCSTKSKAEP-----YVLV 349
Db 906 ----CPRSPGLQBPAPSSYDPLALSELPLGLF---DCEMLDAVDPOHNGYVLV 949

RESULT 3
US-11-072-512-3542
; Sequence 3542, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAWATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 3542
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-072-512-3542

Query Match 5.9%; Score 137; DB 7; Length 143;
Best Local Similarity 25.5%; Pred. No. 0.00027;
Matches 39; Conservative 23; Mismatches 57; Indels 34; Gaps 5;

QY 285 PYENTNSEKPDPTKPDKDVQWQNEWYIGEYSRQAVEDVLMKENK----- 328
Db 5 PLPRTSVVPRPTTAQOE-----TRNGTADAASKGKSGSLPSVAPTGSASAA 51
QY 329 -DQTFVLVRDCSTKSKAEPYVLVVFYGNKVNVKIRFLESNQCFALG-TGLRGNEMFDSVE 386
Db 52 EDGAYTVRSGSGPHGQPTFLAVLLRGRVNTFIRLDGGRHYALGREGNRRELFSVA 111
QY 387 DIIHYTYFPILLIDGDKAARRKQCYL--TOP 417
Db 112 AMQVHFMMHPLPLVD-RHSGSRELTCLLPFTKP 143

RESULT 4
US-10-467-033-2
; Sequence 2, Application US/10467033
; Publication No. US20060019248A1
; GENERAL INFORMATION:
; APPLICANT: Tiziani, Valdenize
; APPLICANT: Reichenberger, Ernest
; APPLICANT: Ueki, Yasuyoshi
; APPLICANT: Olsen, Bjorn R.
; TITLE OF INVENTION: Mutant SH3-Binding Protein Compositions and Methods
; FILE REFERENCES: H0498.70204US00
```

```
; CURRENT APPLICATION NUMBER: US/10/467,033
; CURRENT FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US 60/266,129
; PRIOR FILING DATE: 2001-02-02
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 561
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-467-033-2

Query Match          5.9%; Score 136.5; DB 6; Length 561;
Best Local Similarity 22.4%; Pred. No. 0.0017;
Matches 94; Conservative 44; Mismatches 138; Indels 143; Gaps 23;

QY 66 NNDYEDPEFOLLKAWPSMKILPARPIQSEYADT-----RYFQDMMEAPLLLPKASVS 119
   |||||:
Db 171 NEDYEHDD-----EDDSYLEPDSPEPGRLEDAIMHPPAYPPP--PVP 210
   |||||:
QY 120 TERQTRDRVMTQLEVDKPTFKDVRQRFGKYTKINKTP-LPPPPPAITL----- 170
   |||||:
Db 211 TPR-----KPAFSDMP-----RAHSFTSGKPGPLPPPPPKHGLPDVGLAAE 252
   |||||:
QY 171 -----PKYQPLP--PAPPEE-SSAYFAPKPTTFEVORGRQR----- 205
   |||||:
Db 253 DSKRDPDLCPRRAEPCRPVATPRMSDPLSTMTPTAGLRKPKCFRESASPSPEPWTGPH 312
   |||||:
QY 206 -SAKDFSRVLGABEESHQTK-----PESSCPSSNQ-----TQKSPPAIAS 246
   |||||:
Db 313 GACSTSSAATATATSRNCDKLSFHLSPRGPTSEPPVPANKPKFKIAEDPPREAA 372
   |||||:
QY 247 SSTYMPGKHSIQARDHTGSMQHCQAQRCQAAASHSPR-----MLP-----YENTNS 291
   |||||:
Db 373 ---MPGLFVPPVAPRPPALK-LVPPEAMARPAVLPRPEKQLPHLQSPDGQSFSP 428
   |||||:
QY 292 EKP-DPTKPD-----EKDVQWNEWYIGYSRQAVEDVLMK-----ENKDGTFVLR 335
   |||||:
Db 429 EKPRQPSQADTGDDSDDEYKVPFNPSVFNVTESCEVERLFKATSPRGEQDGLYCIR 488
   |||||:
QY 336 DCSTKSKAEPYVLVWF--YGNKYVNVKIRFLESNQOQFALGTGLRGNEMFDSVEDIIEHY 392
   |||||:
Db 489 NSSTKSGK---VLVWDETSNKNRYRI--FEKDSKFY----LEGEVLFVSGSMVEHY 538

RESULT 5
US-10-467-033-6
; Sequence 6, Application US/10467033
; Publication No. US20060019248A1
; GENERAL INFORMATION:
; APPLICANT: Tiziani, Valdenize
; APPLICANT: Reichenberger, Ernest
; APPLICANT: Ueki, Yasuyoshi
; APPLICANT: Olsen, Bjorn R.
; TITLE OF INVENTION: Mutant SH3-Binding Protein Compositions and Methods
; FILE REFERENCE: H0498.70204U00
; CURRENT APPLICATION NUMBER: US/10/467,033
; CURRENT FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US 60/266,129
; PRIOR FILING DATE: 2001-02-02
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 561
; TYPE: PRT
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (415)..(415)
; OTHER INFORMATION: Xaa = any amino acid
; OTHER INFORMATION: provided that the mutant domain is not wild type
; FEATURE:
; NAME/KEY: MISC_FEATURE
; OTHER INFORMATION: Subspecies: pertenu (CDC-1 strain)

; LOCATION: (418)..(418)
; OTHER INFORMATION: Xaa = any amino acid
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (420)..(420)
; OTHER INFORMATION: Xaa = any amino acid
; OTHER INFORMATION: provided that the mutant domain is not wild type
US-10-467-033-6

Query Match          5.9%; Score 136.5; DB 6; Length 561;
Best Local Similarity 22.4%; Pred. No. 0.0017;
Matches 94; Conservative 44; Mismatches 138; Indels 143; Gaps 23;

QY 66 NNDYEDPEFOLLKAWPSMKILPARPIQSEYADT-----RYFQDMMEAPLLLPKASVS 119
   |||||:
Db 171 NEDYEHDD-----EDDSYLEPDSPEPGRLEDAIMHPPAYPPP--PVP 210
   |||||:
QY 120 TERQTRDRVMTQLEVDKPTFKDVRQRFGKYTKINKTP-LPPPPPAITL----- 170
   |||||:
Db 211 TPR-----KPAFSDMP-----RAHSFTSGKPGPLPPPPPKHGLPDVGLAAE 252
   |||||:
QY 171 -----PKYQPLP--PAPPEE-SSAYFAPKPTTFEVORGRQR----- 205
   |||||:
Db 253 DSKRDPDLCPRRAEPCRPVATPRMSDPLSTMTPTAGLRKPKCFRESASPSPEPWTGPH 312
   |||||:
QY 206 -SAKDFSRVLGABEESHQTK-----PESSCPSSNQ-----TQKSPPAIAS 246
   |||||:
Db 313 GACSTSSAATATATSRNCDKLSFHLSPRGPTSEPPVPANKPKFKIAEDPPREAA 372
   |||||:
QY 247 SSTYMPGKHSIQARDHTGSMQHCQAQRCQAAASHSPR-----MLP-----YENTNS 291
   |||||:
Db 373 ---MPGLFVPPVAPRPPALK-LVPPEAMARPAVLPRPEKQLPHLQSPDGQSFSP 428
   |||||:
QY 292 EKP-DPTKPD-----EKDVQWNEWYIGYSRQAVEDVLMK-----ENKDGTFVLR 335
   |||||:
Db 429 EKPRQPSQADTGDDSDDEYKVPFNPSVFNVTESCEVERLFKATSPRGEQDGLYCIR 488
   |||||:
QY 336 DCSTKSKAEPYVLVWF--YGNKYVNVKIRFLESNQOQFALGTGLRGNEMFDSVEDIIEHY 392
   |||||:
Db 489 NSSTKSGK---VLVWDETSNKNRYRI--FEKDSKFY----LEGEVLFVSGSMVEHY 538

RESULT 6
US-11-221-263-26
; Sequence 26, Application US/11221263
; Publication No. US20060051823A1
; GENERAL INFORMATION:
; APPLICANT: The United States of America as Represented by the Secretary of the
; APPLICANT: Department of Health and Human Services, Centers for Disease Control and
; APPLICANT: Prevention
; APPLICANT: Liu, Hsi
; APPLICANT: Steiner, Bret
; APPLICANT: Berta, Rodes
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING TREPONEMA PALLIDUM
; FILE REFERENCE: 6395-61666
; CURRENT APPLICATION NUMBER: US/11/221,263
; CURRENT FILING DATE: 2005-09-06
; PRIOR APPLICATION NUMBER: US/10/017,168
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: PCT/US00/16425
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/138,981
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Treponema pallidum
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Subspecies: pertenu (CDC-1 strain)
```

US-11-221-263-26

Query Match 5.5%; Score 128.5; DB 7; Length 393;  
Best Local Similarity 21.3%; Pred. No. 0.0048;  
Matches 78; Conservative 41; Mismatches 154; Indels 93; Gaps 14;  
QY 52 RNLGVGGKCNNDYDEPEQLLKAWPSMKILPARPIQES-----EYADT 99  
DB 63 RTLGTVRGGQTSDG-----LSLASLPSR--VPARPAQRDPLSPPPAGHTVPEYRDT 113  
QY 100 RYFQDMMEAFLPPKASVSTERQTRDVRMTQLEEVDPKTFKDVRSORFKGFKYTKINT 159  
DB 114 VIFDD---PRLVSPLSREGGEREVDV-----PKVVPASEREGG----- 150  
QY 160 PLPPRPAILPKYQPLPAPPESSAYFAPKPTFPFVQVGRQPRQSAKDFSRVL----- 214  
DB 151 -----EREVDVPKVVEPASEREGGEREVEDVPKVVEPASEREGGEREVEDVPKVVEPASE 206  
QY 215 --GAEEESHQTK---PESSCPSSNQNTQKSPPAIASSSYMPGKHSIQARDHTGSMOHC 269  
DB 207 REGGEREVDVPKVVEPASEREGGEREVEDVPKVVEPASEREGGEREVASQHTKQPSH-- 264  
QY 270 AQCQAAASHSPRLPYENTNSBKPDPTKPD--EKDVWQNEWYIGYSRQAVEDVLMKEN 327  
DB 265 -----SVNSGAPNQ--FRNPEGELPF-TLPDLSESEIIVVEEQKGRAHPQVIEGAPRGL 316  
QY 328 KQGTFLVDCSTKSKAEPPYLVVYFG-----NKVYNVKIRP-----L 364  
DB 317 QPGEYYVQIAVFHDAIQVQSIHVHYGYEYPIAVEQDIHEGKVRFTVCVGPVQKDERGAVL 376  
QY 365 ESNQOF 370  
DB 377 ENQRF 382

RESULT 7

US-11-109-156-19  
; Sequence 19, Application US/11109156  
; Publication No. US20050250144A1  
; GENERAL INFORMATION:  
; APPLICANT: Toshio Ota  
; APPLICANT: Takao Isogai  
; APPLICANT: Tetsuo Nishikawa  
; APPLICANT: Koji Hayashi  
; APPLICANT: Kaoru Otsuka  
; APPLICANT: Jun-Ichi Yamamoto  
; APPLICANT: Shizuko Ishii  
; APPLICANT: Tomoyasu Sugiyama  
; APPLICANT: Ai Makamatsu  
; APPLICANT: Keiichi Nagai  
; APPLICANT: Tetsuji Otsuki  
; APPLICANT: Shin-Ichi Funahashi  
; APPLICANT: Chiaki Senoo  
; APPLICANT: Jun-Ichi Nezu  
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN KINASE/PROTEIN  
; FILE REFERENCE: 06501-099002  
; CURRENT APPLICATION NUMBER: US/11/109,156  
; CURRENT FILING DATE: 2005-04-19  
; PRIOR APPLICATION NUMBER: US/10/060,065  
; PRIOR FILING DATE: 2002-01-29  
; PRIOR APPLICATION NUMBER: PCT/JP00/05061  
; PRIOR FILING DATE: 2000-07-28  
; PRIOR APPLICATION NUMBER: US 60/159,590  
; PRIOR FILING DATE: 1999-10-18  
; PRIOR APPLICATION NUMBER: US 60/183,322  
; PRIOR FILING DATE: 2000-02-17  
; PRIOR APPLICATION NUMBER: JP 11-248036  
; PRIOR FILING DATE: 1999-07-29  
; PRIOR APPLICATION NUMBER: JP 2000-118776  
; PRIOR FILING DATE: 2000-01-11  
; PRIOR APPLICATION NUMBER: JP 2000-183767  
; PRIOR FILING DATE: 2000-05-02

; PRIOR APPLICATION NUMBER: JP 2000-241899

; PRIOR FILING DATE: 2000-06-09

; NUMBER OF SEQ ID NOS: 43

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 19

; LENGTH: 724

; TYPE: PRT

; ORGANISM: Homo sapiens

US-11-109-156-19

Query Match 5.5%; Score 128; DB 7; Length 724;  
Best Local Similarity 20.6%; Pred. No. 0.012;  
Matches 78; Conservative 45; Mismatches 135; Indels 120; Gaps 14;  
QY 81 PSMKILPARPIQSEVADTRYFQDMMEAFLLLPPKAS----- 117  
DB 85 PTPKPRPRPLVPAP--GSSKTEADVEQALTLPLDLAEQFAPPDIAPLLKLVAEIEKKG 143  
QY 118 -----VSTERQTRDVRMTQLEEVDPKTFK-----DVSRSQREKGFYTKINTPLPPR 165  
DB 144 LECSTLYRTQSSNLAEQLQLDCDTPSVDLEMDVHVLADAFKRYL-----LDLPN 195  
QY 166 PAITLPPKYOPLPAPPESSAYFAP-----KPTFPFVQVGRQPRQSAKDFSRVLGAEE 218  
DB 196 PVTAAVYSEMISLAEVQSSEYIQLLKILRSPIPHQYWLTLQVLLKHFFKL----- 250  
QY 219 ESHHQTPESSCPSSNQNTQKSPPAIASSSYMPGKHSIQARDHTGSM----- 265  
DB 251 -----SQTSSKNLLNARVLSEIFSPMLFRFSAASSDNTENLIKVIEILISTEMN 299  
QY 266 --QHCFAQRCQAAASHSPRLPYENTNSBKPDPTK-----DEKDVWQNEWYIGY 314  
DB 300 EQPAPA-----LP-----PKPPKPTTVANNGNMNNMSLQNAEWYWGDI 338  
QY 315 SROAVEDVLMKENKGTFLVDCSTKSKAEPPYLVVYFGNKNVKIRPLESNOQFALGT 374  
DB 339 SREEVNEKL-RDTADGTFLVDRDASTKMGD-YTLTKKGN--NKLKIFHRDQKYGFS 394  
QY 375 GLRGNEFMDSVEDIIEHY 392  
DB 395 PL----TFSSVVELINHY 408

RESULT 8

US-11-169-041-167  
; Sequence 167, Application US/11169041  
; Publication No. US20060019284A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES FOR PREDICTING ACTIVITY OF  
; TITLE OF INVENTION: COMPOUNDS THAT INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE  
; TITLE OF INVENTION: KINASES AND/OR PROTEIN TYROSINE KINASE PATHWAYS IN LUNG CANCER  
; TITLE OF INVENTION: CELLS  
; FILE REFERENCE: 10001 NP  
; CURRENT APPLICATION NUMBER: US/11/169,041  
; CURRENT FILING DATE: 2005-06-28  
; PRIOR APPLICATION NUMBER: 60/584,405  
; PRIOR FILING DATE: 2004-06-30  
; NUMBER OF SEQ ID NOS: 527  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 167  
; LENGTH: 863  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-169-041-167

Query Match 5.5%; Score 127.5; DB 7; Length 863;  
Best Local Similarity 22.7%; Pred. No. 0.016;  
Matches 68; Conservative 38; Mismatches 124; Indels 69; Gaps 14;  
QY 32 SLSSAKGRCAVLEPLPDHR---RNLGVFGGSKCNNDYDEPEQLLKAWPSMKILPA 88  
DB 336 SSSLVNGRLHEL--PVPKPRGTGTPKSEGTPAPRK-----DPPWITLVQAEPKKK--- 382



```
QY 89 RPTQSEYADTRYFQDMMEAPLLPPKASVSTERQ--TRDVRMTQLEVEDKPT-FKDVRSQ 146
Db 383 -----PAPLPSSSPPPSQDSROVENGTEVAQPSPTASLESK 422
QY 147 RFKGFYTKINKTLPPLPRAI--TLPKYQPLPAPPESSAYFAPKPT-FPEVQGRQ 204
Db 423 PYNFEEEDKEEEAFAAFLATSPALGH--PESTPKSLHPWYGITPTSSPKTKRPA 480
QY 205 RSKDFRVLGAEEESHQTKPESSCPFS-----SNQNTQKS-----PPAIASSY 249
Db 481 RAPASPLALHASRLSH--SEPPSATPSALSVESLSSESASQTAGAELLEPPAVPKSS 538
QY 250 MPQKH-----SIQARDHTGSMQHCPAQRCQAASHSPMLPYENTNSEKPDPTKP 299
Db 539 EPAVHAPGTGPNVSLSTNSSLASSGELVPRVQMPQASPLAP-RTRGSSGQPAPK 596

RESULT 9
US-10-330-773-627
; Sequence 627, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001300
; CURRENT APPLICATION NUMBER: US/10/330,773
; CURRENT FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 627
; LENGTH: 2344
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-330-773-627

Query Match 5.2%; Score 121.5; DB 6; Length 2344;
Best Local Similarity 19.9%; Pred. No. 0.18;
Matches 81; Conservative 45; Mismatches 151; Indels 131; Gaps 17;

QY 46 PLPDHRLNAGVPGGKCNNDYEDPEFQLLKAWPSMKILPAPPIQSEYADTRYFQDM 105
Db 157 PAKSMWNKQGGQGIQVNS-----OFQOEFFSLQAAGDQEKKEKENDENY---- 203
QY 106 MEAPLLLPK-----ASVSTERQTRDVRMTQLEEVKPTFKDVRSQRFKGF 151
Db 204 GPGSLRPPNVACWRDGGKSNSSDQDEKQLQDSEITATSQNIDILKVKVKRIACGP 263
QY 152 KYTKIN-----KTPLPP-----PRPAITLPKKYQPL--PPAPPESSAY--FAPKPTF 195
Db 264 POAKLNGQQPASQYRAMMPPYMFQYPRMAYPPLHGPMPFPPLSLSEANKSLRGRGPPPSW 323
QY 196 PEVQGRPRQSAK-----DPSRVLG--AEESHQTKPES 228
Db 324 ASEPEREISLASBELKDLKFDNLDADEGAWAGAEVNDYTEQLNFSDDDEQGSTSPKE 383
QY 229 SCPSNNQTKSP-----PAIASSYMPGKHSIQARDHTGS 264
Db 384 S--SSEDTAKTPESTENRKEVDEASTKSSQIPAPQPVTKSPYKGGPPFNQER---GP 438
QY 265 MQHCPAQRCQAASHSP---RMLPYENTNSEKPDPTK---PDEKDVW----- 305
Db 439 SSHLPPPPKLLAQHPPPPPDRQIP----GRQGPFPSPKPPVPDNDENIWKQRRKQOSEISAA 494
QY 306 -----QNEWYIGEYSRQAVEVLMKENKDGTFVLVRCSTKSKAEP 345
Db 495 VERARKREBERMEQRKAAQDSNRSE-KETTQVVQEAEPESGAQP 541

RESULT 10
US-11-221-263-24
; Sequence 24, Application US/11221263
```

```
; Publication No. US20060051823A1
; GENERAL INFORMATION:
; APPLICANT: The United States of America as Represented by the Secretary of the
; APPLICANT: Department of Health and Human Services, Centers for Disease Control and
; APPLICANT: Prevention
; APPLICANT: Liu, Hsi
; APPLICANT: Steiner, Bret
; APPLICANT: Berta, Rodes
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING TREPONEMA PALLIDUM
; FILE REFERENCE: 6395-61666
; CURRENT APPLICATION NUMBER: US/11/221,263
; CURRENT FILING DATE: 2005-09-06
; PRIOR APPLICATION NUMBER: US/10/017,168
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: PCT/US00/16425
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/138,981
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Treponema pallidum
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Subspecies: endemicum (Bosnia strain)
US-11-221-263-24

Query Match 5.1%; Score 117; DB 7; Length 428;
Best Local Similarity 20.4%; Pred. No. 0.047;
Matches 79; Conservative 45; Mismatches 163; Indels 100; Gaps 15;

QY 52 RNLAGVPGGKCNNDYEDPEFQLLKAWPSMKILPAPPIQES-----EYADT 99
Db 63 RTLGTVRGSGTSQDG-----LSLASLPSR--VPAQPAQRDPLSSPPAGHTVPEYRDT 113
QY 100 RYFQD-----MMEAPLLPPPKAS--VSTRQTRDVRMT-----QLEEVDPK 138
Db 114 VIFDDPRLVSPLSREVEDVPKVVPEPASEREGGEREVEDVPKVVPEPASEREGGEREVEDVP 173
QY 139 TFKDVRSQRFKGYTKINKTLPPLPRPAITLPKKYQPLPAPPESSAYFAPKPTFPEV 198
Db 174 KVVPEASEREGG-----EREVEDVPKVVPEPASEREGGEREVEDVPKVVPEPAS 220
QY 199 QRGPQRQSAKDFSRVL-----GABEESHQTK--PSSSCPSNNQTKSPPAIASSS 248
Db 221 EREGGEREVEDVPKVVPEPASEREGGEREVEDVPKVVPEPASEREGGEREVEDVPKVVPEPAS 280
QY 249 YMPGKHSIQARDHTGSMQHCPAQRCQAASHSPMLPYENTNSEKPDPTKPD--EKDVWQ 306
Db 281 EREGGEREVASQHTKOPSH-----SVSNSAPNQ--FRNPEGELPF-TLPDLSESEIVV 330
QY 307 NEWYIGEYSRQAVEVLMKENKDGTFVLVRCSTKSKAEPYVLVVFYG-----NKVYN 358
Db 331 PEFQKGRAPQVPIPEGAPRQPGEEYVQVAVFHDAIQVOSIVHRYGVETPIAVEQDIHE 390
QY 359 VKIRF-----LESNQOF 370
Db 391 GKVRFTVCVQKDERGAVLENFORF 417

RESULT 11
US-10-921-231C-1
; Sequence 1, Application US/10821231C
; Publication No. US20050275837A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Dongman
; APPLICANT: Ben-Amotz, Dor
; APPLICANT: Xie, Yong
; APPLICANT: Davison, Vincent J.
; APPLICANT: Mrozek, Melissa
; APPLICANT: Ortiz, Corabi
```





*This Page Blank (uspto)*